

Result No.	Score	Query Match	Length	DB	ID	Description
1	1789	100.0	342	1	PCT-US98-24857-2	Sequence 2, Appli
2	1759.5	98.4	372	13	US-08-627-356-4163	Sequence 4163, Ap
3	1759.5	98.4	372	20	US-09-611-529-5797	Sequence 5797, Ap
4	1193.5	66.7	350	18	US-09-450-969-6115	Sequence 6115, Ap
5	500	27.9	367	27	US-60-215-161-5966	Sequence 5966, Ap
6	485.5	27.1	366	18	US-09-328-353-5462	Sequence 5462, Ap

[illegible]

Db 61 TTYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKDNSLDTHGDETLRAHGEIEVCVDDE 120
QY 121 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQWITNKKEVKQDVYKLRHRHDA 180
Db 121 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQWITNKKEVKQDVYKLRHRHDA 180
QY 181 VLTGRTVELDDPOYTTRIQQDGKNPIKVLKSGNHFHNOQIYQDESTPIWIYTENPNLT 240
Db 181 VLTGRTVELDDPOYTTRIQQDGKNPIKVLKSGNHFHNOQIYQDESTPIWIYTENPNLT 240
QY 241 SNOTHEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFESIYYIDEFILYYAPKLI 300
Db 241 SNOTHEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFESIYYIDEFILYYAPKLI 300
QY 301 GSGSNYQFYQTNVDIEIPDANQFEIVHSELNQNVLTKRKK 342
Db 301 GSGSNYQFYQTNVDIEIPDANQFEIVHSELNQNVLTKRKK 342

RESULT 2

US-08-827-356-4163
; Sequence 4163 Application US/08827356
; GENERAL INFORMATION:

; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 5574
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Gallopington Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530

; COMPUTER READABLE FORM:
; MEDIUM TYPE: PC-DOS/MS-DOS
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,356
; FILING DATE: 01-APR-1997

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,477
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: 60/016,743
; FILING DATE: 02-MAY-1996
; APPLICATION NUMBER: 60/020,016
; FILING DATE: 14-JUN-1996

; INFORMATION FOR SEQ ID NO: 4163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1...372
US-08-827-356-4163

Query Match 98.4%; Score 1759.5; DB 13; Length 372;
Best Local Similarity 98.8%; Pred. No. 2,4e-174;
Matches 339; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDYATQLANWVGQTCVNPVPGAVVYNEGRIVGIGHLRKGDHAEVQALDMAQQNAEGA 60

Db 30 MDYATQLANWVGQTCVNPVPGAVVYNEGRIVGIGHLRKGDHAEVQALDMAQQNAEGA 89
QY 61 TTYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKDNSLDTHGDETLRAHGEIEVCVDDE 120
Db 90 TTYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKDNSLDTHGDETLRAHGEIEVCVDDE 149
QY 121 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQWITNKKEVKQDVYKLRHRHDA 180
Db 150 RASQLYQDFFKAKAKOLPOITVKVSASLDGKQANDNGSQWITNKKEVKQDVYKLRHRHDA 209
QY 181 VLTGRTVELDDPOYTTRIQQDGKNPIKVLKSGNHFHNOQIYQDESTPIWIYTENPNLT 240
Db 210 VLTGRTVELDDPOYTTRIQQDGKNPIKVLKSGNHFHNOQIYQDESTPIWIYTENPNLT 269
QY 241 SNOTHEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFESIYYIDEFILYYAPKL 299
Db 270 SNOTHEIILYKSCDLTTLILHNLKRGVGTLLVEAGPTTTSFESIYYIDEFILYYAPKL 329
QY 300 GSGSNYQFYQTNVDIEIPDANQFEIVHSELNQNVLTKRKK 342
Db 330 GSGSNYQFYQTNVDIEIPDANQFEIVHSELNQNVLTKRKK 372

RESULT 3

US-09-611-529-5797
; Sequence 5797 Application US/09611529
; GENERAL INFORMATION:

; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963USI

; CURRENT APPLICATION NUMBER: US/09/611,529
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14

; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06

; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,334
; PRIOR FILING DATE: 1998-03-06

; PRIOR APPLICATION NUMBER: US 09/036,221
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,137
; PRIOR FILING DATE: 1998-03-06

; PRIOR APPLICATION NUMBER: US 09/036,082
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,081
; PRIOR FILING DATE: 1998-03-06

; PRIOR APPLICATION NUMBER: US 09/036,079
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06

; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356

```
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 5797
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-611-529-5797

Query Match      98.4%; Score 1759.5; DB 20; Length 372;
Best Local Similarity 98.8%; Pred. No. 2.4e-174;
Matches 339; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 MDYAIQANNVQGTGPNPPVGVAVVYVNEGRIYVIGAHLRKGDHAEVQALDMAQQNAEGA 60
Db 30 MDYAIQANNVQGTGPNPPVGVAVVYVNEGRIYVIGAHLRKGDHAEVQALDMAQQNAEGA 89
QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAKVYATKDNSLDTHGDETLRAHGIIEVCVDDE 120
Db 90 TIYTILEPCSHFGSTPPCVNKIIDCKIAKVYATKDNSLDTHGDETLRAHGIIEVCVDDE 149
QY 121 RASQLYQDFKAKAKOLPOITVKVVSASLDGKQANDNGSQOWITNKEVKODVYKLRHRHDA 180
Db 150 RASQLYQDFKAKAKOLPOITVKVVSASLDGKQANDNGSQOWITNKEVKODVYKLRHRHDA 209
QY 181 VLTGRRTVELDDPYTTRIQQDKNPKIKVILSKSGNIHFNOQIYODESTPIWIYENPNT 240
Db 210 VLTGRRTVELDDPYTTRIQQDKNPKIKVILSKSGNIHFNOQIYODESTPIWIYENPNT 269
QY 241 SNOTHIEIYLKSCDLTTLILHNLKRGVGTLLVEAGPTTISEF-SIYYIDEFILYAPKL 299
Db 270 SNOTHIEIYLKSCDLTTLILHNLKRGVGTLLVEAGPTTISEF-LQSNYIDEFILYAPKL 329
QY 300 IGGSGNYQFYQTNVDIEIPDANQFEIVHSELNQNKLRLKK 342
Db 330 IGGSGNYQFYQTNVDIEIPDANQFEIVHSELNQNKLRLKK 372

RESULT 4
US-09-450-969-6115
; Sequence 6115, Application US/09450969
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: PATH99-09A
; CURRENT APPLICATION NUMBER: US/09/450,969
; CURRENT FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 7544
; SEQ ID NO 6115
; LENGTH: 350
; TYPE: PRT
; ORGANISM: S.epidermidis
US-09-450-969-6115

Query Match      66.7%; Score 1193.5; DB 18; Length 350;
Best Local Similarity 65.0%; Pred. No. 1.9e-115;
Matches 223; Conservative 61; Mismatches 58; Indels 1; Gaps 1;

QY 1 MDYAIQANNVQGTGPNPPVGVAVVYVNEGRIYVIGAHLRKGDHAEVQALDMAQQNAEGA 60
Db 8 MDYAIQANNVQGTGPNPPVGVAVVYVNEGRIYVIGAHLRKGDHAEVQALDMAQQNAEGA 67
QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAKVYATKDNSLDTHGDETLRAHGIIEVCVDDE 120
```

```
Db 68 TIYVLEPCTHHGSTPPCVDKIIIEAGISKVIYAVKDTTLVSKGDEILREAGIEVEFYNE 127
QY 121 RASQLYQDFKAKAKOLPOITVKVVSASLDGKQANDNGSQOWITNKEVKODVYKLRHRHDA 180
Db 128 NAAALYRDEFTAKRNEVPEVTVKVSLSLDGKQATDNESKWIINKEVKEDVYQLRHEDA 187
QY 181 VLTGRRTVELDDPYTTRIQQDKNPKIKVILSKSGNIHFNOQIYODESTPIWIYENPNT 240
Db 188 VITGRRTIEADNPLYTRVPDGHPIRVILSKGQDFNQIIEKDTASEIWIYVITNEKLA 247
QY 241 SNOTHIEIYLKSCDLTTLILHNLKRGVGTLLVEAGPTTISEF-SIYYIDEFILYAPKL 299
Db 248 TNKSFILIIINSNCDDTTILQDLYQRGKLLVEAGPNTISQFLQSKHLNELILYAPKL 307
QY 300 IGGSGNYQFYQTNVDIEIPDANQFEIVHSELNQNKLRLKK 342
Db 308 IGGSGKHQYKTDVIDLPEATQFEIVDSKLINQNLKLRKK 350

RESULT 5
US-60-215-161-5966
; Sequence 5966, Application US/60215161
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)A
; CURRENT APPLICATION NUMBER: US/60/215,161
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 5966
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-60-215-161-5966

Query Match      27.9%; Score 500; DB 27; Length 379;
Best Local Similarity 34.4%; Pred. No. 3.7e-43;
Matches 128; Conservative 67; Mismatches 143; Indels 34; Gaps 9;

QY 1 MDYAIQANNVQGTGPNPPVGVAVVYVNEGRIYVIGAHLRKGDHAEVQALDMAQQNAEGA 60
Db 8 MSRALELAYOGRFTTSPNPNVGCIVVKDQIVGEGPHLRAGEPHAEVHALRMAGERAKGA 67
QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAKVYATKDNS--LDTHGDETLRAHGIIEVCV 117
Db 68 TAYVTLEPCSHHKGTPPCSDALIAAGISRVVAMQDPNFPQVAGRGYKLLQQAQISIEHGV 127
QY 118 DDERASQLYQDFKAKAKOLPOITVKVVSASLDGKQANDNGSQOWITNKEVKODVYKLRHR 177
Db 128 MMEQTEMLNKGFLKRMRTGFPYQLKLGASLDGRTALASGESKWIITSPARQDVQKLRAQ 187
QY 178 HDAVLTGRRTVELDDPYTTRIQQDKNPKIKVILSKSGNIHFNOQIYODESTPIWIYENPNT 224
Db 188 CSAILLSATVADDPDLVVRNDELDAETQAVVYVQAEALQPVRIIVDSKNRITPQHVVQ 247
QY 225 DESTPIWI-YTE-----NPNLTSNQTHTIEIYL-----KSCDLTTLILHNLKRGV 269
Db 248 -QTGCQWLTAHTDLGSDSANSADKNEQWPNDEIKLLPVHGTGVDVLLIMQGLGRQVN 306
QY 270 TLLVEAGPTTISEF-SIYYIDEFILYAPKLIGGSGNYQFYQTNVDIEIPDANQFEIVHS 328
Db 307 SVWAECPFLACALLSLGLVDILILYAPKVLGNSAR-GLFAIPELQKUSDAPEFTLIDV 365
QY 329 ELLNQNVKLTLR 340
```


Db 241 YIATSSDKNKMKLYQNHGCEILSIKKGNHIDLSLMLQHLGNMQIDSLVLEGGSLMWS 300
Qy 282 EFSIYYIDEFLIYAPKLIGS-----GNFYQYQTDNDVIEIPDANQFEIYHSELLNQV 335
Db 301 ALEQOIVDELKIYIAPKIFGSAKFPVGGEGISLPNDAIRLKPFAFSGIGNDYLIESEV 359

RESULT 9
US-08-979-616-2
; Sequence 2, Application US/08979616
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: Novel ribg
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/979,616
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: F50533-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 366 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-979-616-2

Query Match 26.1%; Score 467; DB 14; Length 366;
Best Local Similarity 33.1%; Pred. No. 9.5e-40;
Matches 119; Conservative 65; Mismatches 145; Indels 30; Gaps 9;
Qy 1 MDYAIQIANMVGGOTGVNPPVGVVNEGRIVGIGAHLRKGDKHAEOALDMAQOAE 60
Db 7 MKAIAKLAQAGVYVNPVMAIIVDNHIIQGYHEFFGGPHAEARNALKNCRKSPVGA 66
Qy 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNS--LDTHGDETLRAHGIEV---- 114
Db 67 TLYVTLEPCCHFGTPTPCIDAIIDSGITRVVIGSLDCNPVSGKGVKILENNIQTVTGI 126
Qy 115 ---ECVDDERASQYQDFFAKAKOLPOITVKVSASLDGKQANDNGOSOWITNKVKODV 171
Db 127 LENECLN-----LIKFRKYITHQVYVPMKYAMSDGKIATKTNSKWIIEEARKHV 180
Qy 172 YKLRRHDAVLTGRRTVELDDPQVTTIRIQDGKNPVKVLSKSGNIHFNQOIQY---DEST 228
Db 181 HOLRHYVSAINVGVNTVQDDPLTLCLRECKNPIRICDHLRTPLTSKIVKTANDIKT 240
Qy 229 PIWIYENPN-LTNSQTH-TEIYILKS-----CDLTTLHNLKKGCVGTTLLVEAGPTT-S 281
Db 241 YIATSSDKNKMKLYQNHGCEILSIKKGNHIDLSLMLQHLGNMQIDSLVLEGGSLMWS 300

Qy 282 EFSIYYIDEFLIYAPKLIGS-----GNFYQYQTDNDVIEIPDANQFEIYHSELLNQV 335
Db 301 ALEQOIVDELKIYIAPKIFGSAKFPVGGEGISLPNDAIRLKPFAFSGIGNDYLIESEV 359
RESULT 10
US-09-252-691-7588
; Sequence 7588, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstock et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTE
; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252,691B
; CURRENT FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 7588
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-09-252-691-7588

Query Match 25.9%; Score 464; DB 17; Length 402;
Best Local Similarity 33.7%; Pred. No. 2.3e-39;
Matches 115; Conservative 52; Mismatches 128; Indels 46; Gaps 7;
Qy 1 MDYAIQIANMVGGOTGVNPPVGVVNEGRIVGIGAHLRKGDKHAEOALDMAQOAE 60
Db 42 MARAMKLAQGRGRTTHPNPNVGVIVKDGIEVGGFYRAGEPHAEVHALRMAGEKARGA 101
Qy 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGDETLR-----AHGI 112
Db 102 TAYVTLEPCSHHGRTPPCCREALIAGVSRVVAAMQDPNPQVAGRGYRLQOEGIDVSHGL 161
Qy 113 EVECVDDERASQYQDFFAKAKOLPOITVKVSASLDGKQANDNGOSOWITNKVKODVY 172
Db 162 MM-----QDAEALNKGFLKRMRTGFPFQLKLGASLDGRTAMANGESOWITSPQARRDVQ 216
Qy 173 KLRHRHDAVLTGRRTVELDDPQVTTIRIQD-----GKNPIKVLKSGNIHFN 219
Db 217 RLRAQSHAILTSETVILADDPAMTVRWEELNADTQALYPOENLRQPLRIIDSNQRYTPE 276
Qy 220 QOIYQD-----ESTPIWIY-TENPNLT-SNOTHEIITLYLKSCDLTTLHNLKRGV 268
Db 277 HRIVQOPGETWIARTKEDTREWPGQVRSITVPEHNGHL-----DLVVLMLLKGQOV 328
Qy 269 GTLLVEAGPTTISE-FSIYYIDEFLIYAPKLIGSGNYQF 308
Db 329 NSIWEAGPTLAGALLQAGLVDELIVVYVAPKLLGNDARGLF 369

RESULT 11
PCT-US97-02318-354
; Sequence 354, Application PC/TUS9702318
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Raymond
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 1166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA

Qy	61	TIYTLEPCSHFGSTPPCVNKIIDCKIAKV	91
Db	61	TIYTLEPCSHFGSTPPCVNKIIDCKIAKV	91

RESULT 14
US-09-107-532-5675
; Sequence 5675, Application US/09107532
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7308

Query Match	24.7%	Score 442.5;	DB 16;	Length 387;
Best Local Similarity	33.5%;	Pred. No. 3.7e-37;		
Matches 113;	Conservative	53;	Mismatches 116;	Indels 55;
				Gaps 9;

[illegible]

Search completed: November 14, 2000, 10:09:29
Job time: 167 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:09:29 ; Search time 72.12 Seconds
(without alignments)
185.968 Million cell updates/sec

Title: US-08-978-456-4
Perfect score: 498
Sequence: 1 MDYALQLPNMVGXTGVNPP.....NKIIDCKTAXVVLXNKRQFR 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 866916 seqs, 135474527 residues

Total number of hits satisfying chosen parameters: 866916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA:*

1:	/cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084A_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US084B_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US05_NEW_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
24:	/cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
25:	/cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
26:	/cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
27:	/cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	98.0	99	1	PCT-US97-02318-354 Sequence 354, App
2	488	98.0	99	1	PCT-US98-24857-4 Sequence 4, Appli
3	488	98.0	99	14	US-08-903-470-354 Sequence 354, App
4	451	90.6	342	1	PCT-US98-24857-2 Sequence 2, Appli
5	451	90.6	372	13	US-08-827-356-4163 Sequence 4163, Ap
6	451	90.6	372	20	US-09-611-529-5797 Sequence 5797, Ap

7	347	69.7	18	US-09-450-969-6115 Sequence 6115, Ap
8	232	46.6	379	US-06-215-161-5966 Sequence 5966, Ap
9	224	45.0	402	US-09-252-691-7588 Sequence 7588, Ap
10	224	45.0	407	US-09-614-912-78 Sequence 78, Appli
11	224	45.0	407	US-06-170-906-10 Sequence 10, Appli
12	215	43.2	366	1 PCT-US98-25010-2 Sequence 2, Appli
13	215	43.2	366	14 US-08-979-616-2 Sequence 2, Appli
14	215	43.2	366	19 US-09-583-110-4805 Sequence 4805, Ap
15	213	42.8	363	21 US-06-170-906-14 Sequence 14, Appli
16	210	42.2	366	18 US-09-328-352-5462 Sequence 5462, Ap
17	205	41.2	387	16 US-09-107-532-5675 Sequence 5675, Ap
18	203	40.8	401	18 US-09-369-955-5 Sequence 5, Appli
19	193	38.8	378	26 US-09-438-185-873 Sequence 873, App
20	182.5	36.6	352	19 US-09-540-236-3272 Sequence 3272, Ap
21	182.5	36.6	352	21 US-06-128-476-4795 Sequence 4795, Ap
22	180.5	36.2	410	15 US-09-064-935-15 Sequence 15, Appli
23	165	33.1	138	27 US-06-215-161-5704 Sequence 5704, Ap
24	122	24.5	169	18 US-09-328-352-7583 Sequence 7583, Ap
25	112.5	22.6	171	16 US-09-107-532-4012 Sequence 4012, Ap
26	106.5	21.4	173	16 US-09-134-000-4901 Sequence 4901, Ap
27	102.5	20.6	140	19 US-09-583-110-4436 Sequence 4436, Ap
28	102.5	20.6	155	1 PCT-US97-14436-650 Sequence 650, App
29	102.5	20.6	155	14 US-08-911-503-650 Sequence 650, App
30	102.5	20.6	155	14 US-08-911-503A-650 Sequence 650, App
31	99	19.9	599	21 US-06-170-906-15 Sequence 15, Appli
32	96.5	19.4	173	18 US-09-450-969-5242 Sequence 5242, Ap
33	93.5	18.8	156	13 US-08-827-356-3411 Sequence 3411, Ap
34	93.5	18.8	156	20 US-09-611-529-4977 Sequence 4977, Ap
35	89.5	18.0	187	17 US-09-252-691-6455 Sequence 6455, Ap
36	83	16.7	344	1 PCT-US98-06371-1112 Sequence 1112, Ap
37	83	16.7	344	14 US-08-902-615A-378 Sequence 378, App
38	82	16.5	160	18 US-09-328-352-6593 Sequence 6593, Ap
39	82	16.5	333	14 US-08-993-002A-8132 Sequence 8132, Ap
40	82	16.5	344	14 US-08-993-002A-8133 Sequence 8133, Ap
41	82	16.5	345	14 US-08-993-002A-8134 Sequence 8134, Ap
42	80.5	16.2	168	27 US-06-215-161-7636 Sequence 7636, Ap
43	78	15.7	591	8 US-08-403-768-4 Sequence 4, Appli
44	76	15.3	158	1 PCT-US99-10985-16 Sequence 16, Appli
45	76	15.3	158	3 US-07-928-506-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
PCT-US97-02318-354
; Sequence 354, Application PC/TUS9702318
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Raymond
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 1166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/02318

```

; FILING DATE: 19-FEB-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,888
; FILING DATE: 20-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmli, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50006
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
;
; INFORMATION FOR SEQ ID NO: 354:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US97-02318-354

Query Match 98.0%; Score 488; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 7.4e-56;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHRLRGDKHAEVQALDMAQXNAEGA 60
Db 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHRLRGDKHAEVQALDMAQXNAEGA 60

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQR 99
Db 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQR 99
```

```

RESULT 2
PCT-US98-24857-4
; Sequence 4, Application PC/TUS9824857A
; GENERAL INFORMATION:
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Novel ribg
; FILE REFERENCE: P50444-9
; CURRENT APPLICATION NUMBER: PCT/US98/24857A
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 08/978,456
; EARLIER FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(99)
; OTHER INFORMATION: Xaa = Any Amino Acid
; PCT-US98-24857-4
```

```

Query Match 98.0%; Score 488; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 7.4e-56;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHRLRGDKHAEVQALDMAQXNAEGA 60
Db 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHRLRGDKHAEVQALDMAQXNAEGA 60

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQR 99
Db 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQR 99
```

```

RESULT 3
US-08-903-470-354
; Sequence 354, Application US/08903470
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Raymond
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: Novel Compounds
; NUMBER OF SEQUENCES: 1166
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/903,470
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/011,888
; FILING DATE: 20-FEB-1996
; APPLICATION NUMBER: PCT/US97/02318
; FILING DATE: 19-FEB-1997
; APPLICATION NUMBER: PCT/US97/02547
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmli, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GM50006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 354:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-903-470-354
```

```

Query Match 98.0%; Score 488; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 7.4e-56;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHRLRGDKHAEVQALDMAQXNAEGA 60
Db 1 MDYATQLPNMVGXTCVNPVPGAVVNVNIGRIGVIGAHRLRGDKHAEVQALDMAQXNAEGA 60

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQR 99
Db 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAAXVLXNXRQR 99

RESULT 4
PCT-US98-24857-2
; Sequence 2, Application PC/TUS9824857A
; GENERAL INFORMATION:
```

```
; APPLICANT: SmithKline Beecham Corporation
; TITLE OF INVENTION: Novel ribg
; FILE REFERENCE: P50444-9
; CURRENT APPLICATION NUMBER: PCT/US98/24857A
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 08/978,456
; EARLIER FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
PCT-US98-24857-2

Query Match          90.6%; Score 451; DB 1; Length 342;
Best Local Similarity 95.6%; Pred. No. 2.6e-50;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MDYAIQLPMVQGGTGVNPPVGVVNVNEGRIVGIGAHLRKGDKHAEOALDMAQNAEGA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MDYAIQLANMVQGGTGVNPPVGVVNVNEGRIVGIGAHLRKGDKHAEOALDMAQNAEGA 60

Qy 61 TIYITLPCSHFGSTPPCVNKIIDCKIAVV 91
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKV 91

RESULT 5
US-08-827-356-4163
; Sequence 4163, Application US/08827356
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 5574
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,356
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,477
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: 60/016,743
; FILING DATE: 02-MAY-1996
; APPLICATION NUMBER: 60/020,016
; FILING DATE: 14-JUN-1996
; INFORMATION FOR SEQ ID NO: 4163:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc_feature
;
```

```
; LOCATION: 1...372
US-08-827-356-4163

Query Match          90.6%; Score 451; DB 13; Length 372;
Best Local Similarity 95.6%; Pred. No. 2.9e-50;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MDYAIQLPMVQGGTGVNPPVGVVNVNEGRIVGIGAHLRKGDKHAEOALDMAQNAEGA 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 30 MDYAIQLANMVQGGTGVNPPVGVVNVNEGRIVGIGAHLRKGDKHAEOALDMAQNAEGA 89

Qy 61 TIYITLPCSHFGSTPPCVNKIIDCKIAVV 91
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 90 TIYITLPCSHFGSTPPCVNKIIDCKIAKV 120

RESULT 6
US-09-611-529-5797
; Sequence 5797, Application US/09611529
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US1
; CURRENT APPLICATION NUMBER: US/09/611,529
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,334
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,221
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,137
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,082
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,081
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,079
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
;
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:06:42 ; Search time 28.37 Seconds
(without alignments)
412.206 Million cell updates/sec

Title: US-08-978-456-2

Perfect score: 1789

Sequence: 1 MDYAIQLANMVGQGTGNPP.....FEIVHSLLNONVKLTLRKK 342

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36.*
1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	100.0	342	20 Y15913	Pyrimidine deaminase
2	547	30.6	361	12 R10070	B. subtilis rib ORF
3	547	30.6	1855	20 Y21803	B. subtilis rib op
4	547	30.6	1855	21 Y83271	Polypeptide encode
5	467	26.1	366	20 Y17949	S. pneumoniae ribg
6	451	25.2	99	18 W28084	Amino acid sequenc
7	451	25.2	99	20 Y15914	Pyrimidine deamina
8	435	24.3	376	20 Y35199	Chlamydia pneumoni
9	393	22.0	410	19 W55031	Actinobacillus ple
10	393	22.0	410	20 Y21972	App ribg gene prod
11	356	19.9	401	21 Y96284	Sinorhizobium meli
12	277	15.5	249	20 W81550	Bacillus subtilis

13	234.5	13.1	261	20	Y37427	Amino acid sequenc
14	194	10.8	170	20	Y37428	Amino acid sequenc
15	153	8.6	155	19	Y86119	S. pneumoniae deri
16	140	7.8	344	19	W98756	H. pylori GHPO 107
17	138.5	7.7	373	20	Y53913	Amino acid sequenc
18	128.5	7.2	158	11	R08285	Thermally stable c
19	128.5	7.2	158	16	R68507	Yeast cytosine-dea
20	127.5	7.1	158	21	Y55104	Chimeric cytosine
21	125.5	7.0	158	21	Y55103	Chimeric cytosine
22	122.5	6.8	158	21	Y55101	Chimeric cytosine
23	121.5	6.8	246	16	R88283	Ashbya gossypii HT
24	119.5	6.7	158	21	Y55102	Chimeric cytosine
25	117.5	6.6	609	16	R88279	Ashbya gossypii DR
26	117	6.5	244	15	R52829	HTP-reductase. Sa
27	111	6.2	163	20	Y35657	Chlamydia pneumoni
28	107.5	6.0	176	20	Y37278	Amino acid sequenc
29	104	5.8	591	15	R52825	DRAP-deaminase. S
30	101.5	5.7	245	18	W89774	Staphylococcus aur
31	98	5.5	1277	21	Y91928	Paenibacillus pabu
32	93	5.2	670	15	R65795	APP751 beta-amyloi
33	93	5.2	770	10	P94775	Novel amyloid prec
34	93	5.2	770	13	R26340	APP770. Homo sapi
35	93	5.2	770	14	R41546	Mutated APP770 exo
36	93	5.2	770	15	R63442	Amyloid protein pr
37	93	5.2	770	15	R62505	Amyloid precursor
38	93	5.2	770	18	W19500	APP770 mutant A-be
39	93	5.2	770	18	W19488	APP770 mutant A-be
40	93	5.2	770	18	W19503	APP770 mutant A-be
41	93	5.2	770	18	W19506	APP770 mutant A-be
42	93	5.2	770	18	W19497	APP770 mutant A-be
43	93	5.2	770	18	W19491	APP770 mutant A-be
44	93	5.2	770	18	W19485	APP770 mutant A-be
45	93	5.2	770	18	W19482	APP770 mutant A-be

ALIGNMENTS

RESULT 1
Y15913
ID Y15913 standard; Protein; 342 AA.
XX
AC Y15913;
XX
DT 04-AUG-1999 (first entry)
DE Pyrimidine deaminase and pyrimidine reductase (ribG) polypeptide.
XX
KW Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;
KW agonist; antimicrobial; antibiotic; Helicobacter pylori infection;
KW H. pylori-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KW gastritis.
XX
OS Staphylococcus aureus.
XX
PN WO926475-A1.
XX
PD 03-JUN-1999.
XX
PF 23-NOV-1998; 98WO-US24857.
XX
PR 25-NOV-1997; 97US-0978456.
XX
(SMIK) SMITHKLINE BEECHAM CORP.
PI Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas RO, Palmer LM, Pratt JM, Reichard RW;
PI Rosenberg M, Trani CM, Ward JM, Warren RL;
XX WPI; 1999-347572/29.
DR N-PSDB; X59909.
XX
PT New Staphylococcus aureus polypeptide and polynucleotide useful in

PT the treatment of gastric ulcer and gastritis

XX
PS Claim 8; Page 5; 48pp; English.

XX
CC The present sequence represents a *Staphylococcus aureus* pyrimidine
CC deaminase and pyrimidine reductase (ribG) polypeptide. *Staphylococcus*
CC *aureus* ribG and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribG can
CC be determined by analysing the nucleic acid sequence encoding ribG
CC or detecting the ribG polypeptide in a sample. ribG can also be used
CC to identify antagonists or agonists. ribG, or its related nucleic acid,
CC also has use as a vaccine to induce an immunological response in an
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
CC especially broad-spectrum antibiotics, may be of use in the treatment
CC of *Helicobacter pylori* infection. This should decrease the advent of
CC *H. pylori*-induced cancers, such as gastrointestinal carcinoma. The
CC treatment should also cure gastric ulcers and gastritis.

XX
SQ Sequence 342 AA;

Query Match 100.0%; Score 1789; DB 20; Length 342;

Best Local Similarity 100.0%; Pred. No. 1.1e-159; Indels 0; Gaps 0;

Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDYAIQIANNVQGTGPNPPVGVAVVNEGRIVGIGAHLRKDKHAEVQALDMAQQNAEGA 60
DB 1 mdyaiaqlannvqgtgpnppvgavvnegrivgigahlrkdkhaevqaldmaqqnaega 60
QY 61 TIYTILEPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGDETLRAHGIEVCVDDE 120
DB 61 tiyitlepcshfgstppcvnkkiiidckiakvvyatknsldthgdetlrahgievecvdde 120
QY 121 RASQYQDFKAKAKOLPQITVKVSASLDGKQANDNGSQWITNKEVKQDVYKLRHRHDA 180
DB 121 rasqlyqdfkakaqlpqitvkvsasldgkqandngsqwitinkevkdvyklrhrhda 180
QY 181 VLTCRRRTVELDDPQYTTTRIQDGNPKIVILSKSGNIHFNQIYQDESTPIWIYTENPNT 240
DB 181 vltrrrtvelddpqytrtiqdgknpiivilsksgnihfnqiyqdestpiwiytentpnt 240
QY 241 SNQTHIEIYKSCDLTTLIHLNLYKRGVGTLLVEAGPTTTSFPIYIDFILIYAPKLI 300
DB 241 sndthieilykscdlttlihllykrvgvtllveagptttsfpiyidefiliyapkli 300
QY 301 GSGNGYQYQTNVDIETPDANQFEIVHSELLNQNVKLTLRKK 342
DB 301 gsgngyqyqtnvdielpdanqfeivhsellnqnkvltlrkk 342

RESULT 2

ID R10070 standard; Protein; 361 AA.

XX
AC R10070;

XX
DT 19-MAR-1991 (first entry)

XX
DE B. subtilis rib ORF 5 gene product.

XX
KW Riboflavin; vitamin B2.

XX
OS Bacillus subtilis.

XX
PN EP405370-A.

XX
PD 02-JAN-1991.

XX
PF 22-JUN-1990; 90EP-0111916.

XX
PR 22-JUN-1989; 89US-0370378.

XX
PA (HOFF) HOFFMANN-LA ROCHE AG.

XX Perkins JB, Pero JG, Sloma A;

XX WPI: 1991-008756/02.

XX Q-PSDB; Q10165.

XX DNA encoding riboflavin biosynthetic proteins - used to produce
PT recombinant bacteria for prodn. of riboflavin

XX Disclosure; Fig 3; 72pp; English.

XX The peptide is a product of the B. subtilin rib operon, involved in
CC the biosynthesis of riboflavin.
CC The product may be expressed in a transformed bacterial expression
CC system for the efficient production of riboflavin (vitamin B2).

XX
SQ Sequence 361 AA;

Query Match 30.6%; Score 547; DB 12; Length 361;

Best Local Similarity 36.1%; Pred. No. 3.2e-43;

Matches 130; Conservative 68; Mismatches 134; Indels 28; Gaps 8;

QY 1 MDYAIQIANNVQGTGPNPPVGVAVVNEGRIVGIGAHLRKDKHAEVQALDMAQQNAEGA 60
DB 6 mklaldldakggegtesnplvgavvkdqgvlgmghlkygeahaevhahmagahaega 65
QY 61 TIYTILEPCSHFGSTPPCVNKKIIDCKIAKVYATKNS--LDTHGDETLRAHGIEV--ECV 117
DB 66 diytilepcshygtktpcaellinsgkrvfamrdpnplvagrismmkeaglevregi 125
QY 118 DDFRASOLYQDFKAKAKOLPQITVKVSASLDGKQANDNGSQWITNKEVKQDVYKLRHR 177
DB 126 ladgaerlnekflhfmrtglpyvtlkaasldgklatstgdskwitseaarqdaqyrtk 185
QY 178 HDVAVLGRRTVELDDPQYTTTRIQD-GKNPKIVILSKSGNIHFNQIYQDESTPIWIY-- 234
DB 186 hqsilyvgvtvakdpsaltcrilpnvtkqpvrviltdvlsipedakvicddqaptwifta 245
QY 235 -----ENPNLTSNOTHIEIYKSCDLTTLIHLNLYKRGVGTLLVEAGPTTTSF--SIYI 288
DB 246 radeeekkrisafgnvftteteriqipdvklilaecgimsyveggsvahgsfvkegcf 305
QY 289 DEFILYAPKLIIGG-----SG-NYQYQTNVDIETPDANQFEIVHSELLNQNVKLTLR 340
DB 306 qeilyfapkligthapslisgegfgsmkdvpilqftditq-----igrdikltak 357

RESULT 3

ID Y21803 standard; Protein; 1855 AA.

XX
AC Y21803;

XX
DT 10-SEP-1999 (first entry)

XX
DE B. subtilis rib operon protein translated from reading frame 3.

XX
KW Riboflavin; open reading frame; ORF; structural gene; promoter;
KW vitamin B2; Bacillus subtilis; rib operon.

XX
OS Bacillus subtilis.

XX
FH Key Location/Qualifiers

FT Misc-difference 1..1855 /note= "residues Xaa are encoded by stop codons"

FT US592538-A.

XX
PN 20-JUL-1999.

XX
PD 24-AUG-1998; 98US-0138775.

XX

CC a sample. ribG can also be used to identify antagonists or agonists.
CC ribG, or its related nucleic acid, also has use as a vaccine to induce
CC an immunological response in an animal. Antimicrobial compounds (e.g.
CC agonists and antagonists of ribG), especially broad-spectrum antibiotics,
CC may be of use in the treatment of *Helicobacter pylori* infection. This
CC should decrease the advent of *H. pylori*-induced cancers, such as
CC gastrointestinal carcinoma. The treatment should also cure gastric ulcers
CC and gastritis. The present sequence represents the *S. pneumoniae* ribG
CC polypeptide.
xx
SQ Sequence 366 AA;

Query Match 26.1%; Score 467; DB 20; Length 366;
Best Local Similarity 33.1%; Pred. No. 1e-35;
Matches 119; Conservative 65; Mismatches 145; Indels 30; Gaps 9;
QY 1 MDYATOLANVQGTGVNPPGAVVNEGRIVGTGAHLRKGDKHAHQALDMAQQAEGA 60
Db 7 mklaklaqkagayvnpmpvgalivdnhilgqgyheffgpphaernalknckrspvga 66
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETLRAHGIEV--- 114
Db 67 tlyvtlpcchfgktpccidaiidsgitrvvigslcdnpivgkvgkileennlqvtvgi 126
QY 115 ---ECVDDERASQLYQDFFKAKAKOLPQITVVKVSASLDGKQANDNGOSQWTINKEVKQDV 171
Db 127 lenecln-----lksfrkyitghvpyvmkymamsmdgkiatknqskwiteearkhv 180
QY 172 YKLRRHDAVLTGRVTVELDPOVTTTQDCKNPIKVLKSGNIHFNQIYQ---DEST 228
Db 181 hqlrhysalnvgtvtdqdpdlitclreegknpirlicdthlrtpitkskvktandikt 240
QY 229 PIWYVTENPN-LTSNQVH-IEIIVLKS----CDLTTLNLNLYKRGVGTLLVEAGPTTT-S 281
Db 241 ylatssedknmklyqhngceilsikkgnhidlsslmqhlgnmqidslvlegslmws 300
QY 282 EFSYIIDEFLYAPKLIGGS-----GNVQFYQTNVDVIEIPDANQPEIVHSELLNQNV 335
Db 301 aleqgidvelkiyapklfggsakfpvggegislpndairikpyafsqigndyliesev 359

RESULT 6

W28084
ID W28084 standard; Protein; 99 AA.

AC W28084;

DT 01-SEP-1998 (first entry)

DE Amino acid sequence of riboflavin biosynthesis protein RIBG.

XX Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome.
XX Staphylococcus aureus.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

FT Misc-difference 1..99
FT /note= "residues designated X are not defined in
FT the specification"

XX W09730070-A1.

XX 21-AUG-1997.

XX 19-FEB-1997; 97WO-US02318.

XX 20-FEB-1996; 96US-0011888.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX

PI Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;

PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;

DR WPI; 1997-424969/39.

DR N-PSDB; T84023.

XX Novel polypeptide(s) from *Staphylococcus aureus* strain WCUH29 - used
PT to isolate antimicrobial compounds, and in vaccines against *S.*
PT aureus infection

XX Claim 6; Page 454-455; 989pp; English.

XX The present sequence represents a *Staphylococcus aureus* protein, that,
CC based on homology with a *Bacillus subtilis* protein, is believed to be a
CC riboflavin biosynthesis protein RIBG. The DNA sequence was isolated from
CC a library of clones of *S. aureus* WCUH 29 in *Escherichia coli*. The DNA
CC sequence can be used in the construction of ribozymes and antisense
CC sequences to control the expression of staphylococcal genes. The DNA
CC sequence is also useful as a source of regulatory elements for the
CC control of bacterial gene expression. The present protein may be used
CC to produce vaccines to enable a host to produce specific antibodies
CC with antibacterial action. These vaccines and antibodies would protect
CC a host against invasion by *S. aureus*, and conditions relating to
CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
CC skin syndrome, and toxic shock syndrome.

XX Sequence 99 AA;

Query Match 25.2%; Score 451; DB 18; Length 99;
Best Local Similarity 95.6%; Pred. No. 4.8e-35;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDYATOLANVQGTGVNPPGAVVNEGRIVGTGAHLRKGDKHAHQALDMAQQAEGA 60

Db 1 mdyaqlpnmvgnvgtvnpvgavvnegrivgahlrkgdkhaevgaldmaqnaga 60

QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAKV 91

Db 61 tiyitlpcshfgstppcvnkiidckiaxvv 91

RESULT 7

Y15914
ID Y15914 standard; Protein; 99 AA.

AC Y15914;

DT 04-AUG-1999 (first entry)

DE Pyrimidine deaminase and pyrimidine reductase (ribG) encoded by ORF.

XX Pyrimidine deaminase; pyrimidine reductase; ribG; antagonist;
KW agonist; antimicrobial; antibiotic; *Helicobacter pylori* infection;
KW *H. pylori*-induced cancer; gastrointestinal carcinoma; gastric ulcer;
KW gastritis.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

FT Misc-difference 1..99
FT /note= "X= any amino acid or stop codon"

XX W09926475-A1.

XX 03-JUN-1999.

XX 23-NOV-1998; 98WO-US24857.

XX 25-NOV-1997; 97US-0978456.

```

XX Black MT, Burnham MKR, Fedon JC, Hodgson JE, Knowles DJC;
PI Lonetto MA, Nicholas RO, Palmer IM, Pratt JM, Reichard RW;
PI Rosenberg M, Trani CM, Ward JM, Warren RL;
XX
XX WPI: 1999-347572/29.
DR N-PSDB; X59910.
XX
XX New Staphylococcus aureus polypeptide and polynucleotide useful in
PT the treatment of gastric ulcer and gastritis
XX
XX Claim 23; Page 6; 48pp; English.
XX
XX The present sequence is encoded by the ORF of Staphylococcus aureus
CC pyrimidine deaminase and pyrimidine reductase (ribG) DNA. Staphylococcus
CC aureus ribG and its antagonists are used to treat individuals in
CC need of them. Disease related to expression or activity of ribG can
CC be determined by analysing the nucleic acid sequence encoding ribG
CC or detecting the ribG polypeptide in a sample. ribG can also be used
CC to identify antagonists or agonists. ribG, or its related nucleic acid,
CC also has use as a vaccine to induce an immunological response in an
CC animal. Antimicrobial compounds (e.g. agonists and antagonists of ribG),
CC especially broad-spectrum antibiotics, may be of use in the treatment
CC of Helicobacter pylori infection. This should decrease the advent of
CC H. pylori-induced cancers, such as gastrointestinal carcinoma. The
CC treatment should also cure gastric ulcers and gastritis.
XX
XX Sequence 99 AA;
SQ
Query Match 25.2%; Score 451; DB 20; Length 99;
Best Local Similarity 95.6%; Pred. No. 4.8e-35;
Matches 87; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 MDYAIQLANMVQGTGVNPPVGAVVYVNEGRIYVIGIGAHLRKGDHAEVQALDMAQNAEGA 60
Db 1 mdyaiglpmmvgxtgvnppvgavvynegrivgigahlrkgdkhaevqaldmaqnaega 60
QY 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKV 91
Db 61 tiyitlpcshfgstppcvnkiidckiaxv 91
RESULT 8
Y35519
ID Y35519 standard; Protein; 376 AA.
XX
AC Y35519;
XX
XX 13-SEP-1999 (first entry)
XX
XX Chlamydia pneumoniae transmembrane protein sequence.
XX
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX
XX Chlamydia pneumoniae.
OS
XX WO9927105-A2.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-IB01890.
XX
XX 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
XX (CEST ) GENSET.
XX
XX Griffais R;
PI
XX WPI: 1999-357842/30.
DR

```

```

XX Genome sequence of Chlamydia pneumoniae
PT
XX Page 1273-1274; Disclosure; 1912pp; English.
XX
XX Y34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
XX Sequence 376 AA;
SQ
Query Match 24.3%; Score 435; DB 20; Length 376;
Best Local Similarity 31.2%; Pred. No. 1.1e-32;
Matches 112; Conservative 65; Mismatches 152; Indels 30; Gaps 9;
QY 1 MDYAIQLANMVQGTGVNPPVGAVVYVNEGRIYVIGIGAHLRKGDHAEVQALDMAQNAEGA 60
Db 12 mrraieigekgritappnpwgcgvvqenriiegfhayagggphaeeaiqnasmpisgs 71
QY 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKVYA--TKDNSLDTHGDETURAHGIEVEC-V 117
Db 72 dvvyslepcshfgstppcvnkiidckiaxvrfvalvdppkavagqgiarqagivvygi 131
QY 118 DDERASQLYQDFEFAKAKOLPQITVKVSASLDGKOANDNGSQWITNKVKQDVYKLRHR 177
Db 132 geseqaslpqlyqrthnfpwtiklsadvgqvdsgqskwtccpearhadvkrlae 191
QY 178 HDAVLTGRTVELDPQYTTTRIQQG----KNPIKVLTSKSGNIHFNQIYODESTPIWY 233
Db 192 sqailvgstrvlsddpwtlcarpqgmlypkqlrvldsrsgvppskvfdktsptlyvt 251
QY 234 TENPNLTSNQTHIEIYIYLSCDLFTI-----LHNLV-----KRGVGTLLVEAGPTT 279
Db 252 ter----cpenyikv--ldslavpvltestpsgvdhkhkvyeylaqkklqvlvegggtl 305
QY 280 -TSEFSIYVIDEFLYAPKLIGSGNYQFYQTNVDVIEIPDANQPFIVHSELLNONVKL 337
Db 306 htsllkerfvnslvlysgpmiigdkrplvgvlgnlle--saspltkssqilgnsikv 362
RESULT 9
W55031
ID W55031 standard; Protein; 410 AA.
XX
AC W55031;
XX
XX 12-OCT-1998 (first entry)
XX
XX Actinobacillus pleuropneumoniae rib-specific deaminase.
XX
XX Vaccine; porcine pleuropneumonia; attenuation; riboflavin;
KW vitamin B2; ribGBAH operon; RibG protein; deaminase;
KW Pasteurella multocida; Pasteurella haemolytica;
KW Haemophilus parausis; Haemophilus influenzae; Haemophilus ducreyi;
KW bronchopneumonia; fibrinous pleuritis.
XX
XX Actinobacillus pleuropneumoniae ISU178 (serotype 5).
OS
XX WO9818917-A2.
XX
XX 07-MAY-1998.
XX
XX 28-OCT-1997; 97WO-US19649.
XX
XX

```



```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-104-165-3

Query Match          5.2%; Score 93; DB 2; Length 770;
Best Local Similarity 21.5%; Pred. No. 0.28;
Matches 66; Conservative 41; Mismatches 96; Indels 104; Gaps 14;

QY 12 QGOTGVNPPVGA---VNEGRIV-----GIGAHLRKGDKHAQVQALDMAQQAEGA 60
   | : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 294 QAETG---PCRAMISRWYFDVTEGKCAPFYGCGGNRRNFD--TEECMAVCGSAMSO 348
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 61 TYITILEPCSH-----FGSTPPCVNKIIDCKIAKVYVATKDNSLDTHGDETLRAH-- 110
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 349 LKTTQEPPLARDVPKLPPTTAASTPDAVDKY-----LETPGDENEHAHFQ 392
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 111 --GIEVECDVDERASQYQDFKA--KAKOLPOITVKV-----SASLDGKQANDNGQ 158
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 393 KAKERLEAKHRMSQVMREWEAEARQAKNLPKADKKAVIQHFQEKVESLEQEAAN---- 448
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 159 SQWITNKEVKQDVYKLRHRHDAVLTGRTTVELDDPOYTTRIQ----- 200
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 449 -----ERQQLVETHMARVEAMLNDRRRLALEN--YITALQAVPPRPRHVNMLKKYVR 499
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 201 -----DGKNPIKVLKSGNIHFNQIYQDESTPIWIYTENPNLTSNOTHIEIYLSKCDL 256
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 500 AEQKDRQHTLK-----HFEHVRMVDPKKAAQIRSQ-----VMTHLRVIYIERMNQS 544
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 257 TTILHNL 263
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 545 LSLLYNV 551

RESULT 12
US-08-464-250-3
; Sequence 3, Application US/08464250
; Patent No. 6107542
; GENERAL INFORMATION:
; APPLICANT: HARDY, John Anthony
; APPLICANT: GOATE, Alison Mary
; APPLICANT: MULLAN, Michael John
; APPLICANT: CHARTIER-HARLIN, Marie-Christine
; TITLE OF INVENTION: Test and Model for Alzheimer's Disease
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,250
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/104,165
; FILING DATE: 21-JAN-1992
; APPLICATION NUMBER: 9101307.8
; FILING DATE: 21-JAN-1991
; APPLICATION NUMBER: 9118445.7
; FILING DATE: 28-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 16163-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-464-250-3

Query Match          5.2%; Score 93; DB 3; Length 770;
Best Local Similarity 21.5%; Pred. No. 0.28;
Matches 66; Conservative 41; Mismatches 96; Indels 104; Gaps 14;

QY 12 QGOTGVNPPVGA---VNEGRIV-----GIGAHLRKGDKHAQVQALDMAQQAEGA 60
   | : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 294 QAETG---PCRAMISRWYFDVTEGKCAPFYGCGGNRRNFD--TEECMAVCGSAMSO 348
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 61 TYITILEPCSH-----FGSTPPCVNKIIDCKIAKVYVATKDNSLDTHGDETLRAH-- 110
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 349 LKTTQEPPLARDVPKLPPTTAASTPDAVDKY-----LETPGDENEHAHFQ 392
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 111 --GIEVECDVDERASQYQDFKA--KAKOLPOITVKV-----SASLDGKQANDNGQ 158
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 393 KAKERLEAKHRMSQVMREWEAEARQAKNLPKADKKAVIQHFQEKVESLEQEAAN---- 448
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 159 SQWITNKEVKQDVYKLRHRHDAVLTGRTTVELDDPOYTTRIQ----- 200
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 449 -----ERQQLVETHMARVEAMLNDRRRLALEN--YITALQAVPPRPRHVNMLKKYVR 499
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 201 -----DGKNPIKVLKSGNIHFNQIYQDESTPIWIYTENPNLTSNOTHIEIYLSKCDL 256
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 500 AEQKDRQHTLK-----HFEHVRMVDPKKAAQIRSQ-----VMTHLRVIYIERMNQS 544
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
QY 257 TTILHNL 263
   : | | | : | | : | | : | | : | | : | | : | | : | | : | | :
Db 545 LSLLYNV 551

RESULT 13
US-08-123-702-4
; Sequence 4, Application US/08123702
; Patent No. 5604131
; GENERAL INFORMATION:
; APPLICANT: Wadsworth, Samuel
; APPLICANT: Snyder, Benjamin
; APPLICANT: Reddy, Vermuri, B.
; APPLICANT: Wei, Chamer
; TITLE OF INVENTION: A cDNA Genomic Hybrid Sequence Encoding APP770
; Patent No. 5604131
; TITLE OF INVENTION: Containing a Genomic DNA Insert of the KI and OX-2 Regions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
```

```

1  COUNTRY: US
2  ZIP: 94301
3
4  COMPUTER READABLE FORM:
5
6  MEDIUM TYPE: Floppy Disk
7
8  COMPUTER: IBM PC compatible
9
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11
12 CURRENT APPLICATION DATA:
13
14 APPLICATION NUMBER: US/08/104,165
15
16 FILING DATE: 21-JAN-1992
17
18 CLASSIFICATION: 435
19
20 PRIOR APPLICATION DATA:
21
22 APPLICATION NUMBER: 9101307.8
23
24 FILING DATE: 21-JAN-1991
25
26 APPLICATION NUMBER: 9118445.7
27
28 FILING DATE: 28-AUG-1991
29
30 ATTORNEY/AGENT INFORMATION:
31
32 NAME: Liebeschuetz, Joe
33
34 REGISTRATION NUMBER: 37,505
35
36 REFERENCE/DOCKET NUMBER: 16163-000100
37
38 TELECOMMUNICATION INFORMATION:
39
40 TELEPHONE: (415) 326-2400
41
42 TELEFAX: (415) 326-2422
43
44 INFORMATION FOR SEQ ID NO: 2:
45
46 SEQUENCE CHARACTERISTICS:
47
48 LENGTH: 751 amino acids
49
50 TYPE: amino acid
51
52 STRANDEDNESS: single
53
54 TOPOLOGY: linear
55
56 MOLECULE TYPE: protein
57
58 US-08-104-165-2

```

Query Match	5.1%;	Score 91.5;	DB 2;	Length 751;
Best Local Similarity	22.1%;	Pred. No. 0.39;		
Matches	66;	Conservative	38;	Mismatches 88;
				Indels 107;
				Gaps 14;

Qy	12	QGOTGVNPDVGAVV-----VNEGRIV-----GICAHLRKGDHAEVQALDMAQAQNAEGA	60
Db	294	QAETG--PCRAMISWYEDVTGECAPFYGGCGGNRNFTTEEYCMAVC-----GS	343
Qy	61	TYITILEPCHSGFSTPPCKNIIDCKIAKVVYATKDNSLDTHGDETLEAH----GIEVEBC	116
Db	344	AIPPTA-----ASTPDAVDKY-----LETPGDENEHAHFQAKERLEA	381
Qy	117	VDDERASQLYQDFFA--KAKOLPIQTVKV-----SASLDGQANDNGOSQMITNKE	166
Db	382	KHRMSQVMREWEAEARQAKNLPKADKKAVIQHFQEKVESLEQEAAAN-----E	430
Qy	167	VHQDVKLVRHRHDVAVLTRGTVELDDPOVTTTIIQ-----DGKN	204
Db	431	RQOVLVETHMARVEAMLNDERRALEN--YITALQAVPPRPHRVFNWLKKYYRAEQKDROH	488
Qy	205	PKIVILSKSGNIHFNOOIYQDESTPIWIIVTENPNLTISNOTHEIYLKSCDLTTILHNLI	263
Db	489	TLLK-----HFBEHVRMVDPKKAAQIRSQ-----VMTHLURVIYERMNQSLSYLVN	532

RESULT 15
US-08-422-333-2
; Sequence 2, Application US/08422333
; Patent No. 5912410
; GENERAL INFORMATION:
; APPLICANT: CORDELL, Barbara L.
; TITLE OF INVENTION: TRANSGENIC NON-HUMAN MAMMAL DISPLAYING
; TITLE OF INVENTION: THE AMYLOID-FORMING PATHOLOGY OF ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scios, Inc.
; STREET: 2450 Bayshore Parkway
; CITY: Mountain View
; STATE: CA
; COUNTRY: USA
; ZIP: 94043

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/422,333
; FILING DATE: 13-APR-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Shearer, Peter R.
; REGISTRATION NUMBER: 28,117
; REFERENCE/DOCKET NUMBER: 21900-28048.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 966-1550
; TELEFAX: (415) 968-2438
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-422-333-2

Query Match      5.1%; Score 91.5; DB 2; Length 751;
Best Local Similarity 22.1%; Pred. No. 0.39;
Matches 66; Conservative 38; Mismatches 88; Indels 107; Gaps 14;

QY 12 QGOTGVNPPVGAVV-----VNEGRIV-----GIGAHLRKGDKHAEVQALDMAQQNAEGA 60
Db 294 QAETG---PCRAMISRWFYFDVTEGKCAPFFYGGCGGNRNFDTEEYCMAVC-----GS 343
QY 61 TIYTILEPCSHFGSTPCVYNKIIDCKIAKVYATKDNSLDTHGDETLRAH---GIEVEC 116
Db 344 AIPPTA-----ASTPDVDKY-----LETPGDENEHAHFQKAKERLEA 381
QY 117 VDDERASQLYODFFKA--KAKQLPQITVKV-----SASLDGKOANDNGOSQWITNKE 166
Db 382 KHERNSQVMREWEAEFRQAKNLPKADKRAVIQHFQKVESLEQEAAN-----E 430
QY 167 VKODVYKLRHRHDAVLTGRTVELDDPQYTRIQ-----DGNK 204
Db 431 RQQLVETHMARVEAMLNDRRLALEN--YITALQAVPPRPHVFNMLKKYVRAEQKDROH 488
QY 205 PIKVILSKSGNIHFNQIYODESTPIWIYTNPNLTSTNTHIEIYLYKSCDLTTLIHLN 263
Db 489 TLK-----HFEHVRMVDPKAAQIRSQ-----VMTHLRVIYERMNQSLSLDYN 532
```

Search completed: November 14, 2000, 10:07:45
Job time: 63 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:06:42 ; Search time 24.32 Seconds
(without alignments)
892.409 Million cell updates/sec

Title: US-08-978-456-2
Perfect score: 1789
Sequence: 1 MDYAIQLANMVGQGTGVNPP.....FEIVHSELLNONVKLTLRKK 342

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_65:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	547	30.6	361	2 PN0100	riboflavin-specifi
2	545	30.5	356	2 C70313	riboflavin-specifi
3	530	29.6	348	2 G72207	riboflavin-specifi
4	495.5	27.7	372	2 H64103	riboflavin-specifi
5	484	27.1	369	2 B81039	riboflavin-specifi
6	480	26.8	369	2 B81984	probable diaminohy
7	470	26.3	367	2 S26201	riboflavin-specifi
8	435	24.3	376	2 G72026	riboflavin-specifi
9	434.5	24.3	339	2 F70901	probable riboflavi
10	422	23.6	357	2 G75552	riboflavin-specifi
11	413.5	23.1	368	2 S74372	riboflavin-specifi
12	411.5	23.0	396	2 D81472	riboflavin-biosynt
13	408.5	22.8	379	2 E71479	riboflavin-specifi
14	320	17.9	599	2 T12994	probable riboflavi
15	300	16.8	363	2 T10638	riboflavin biosynt
16	227.5	12.7	336	2 D81258	hypothetical prote
17	195	10.9	219	1 F69500	probable riboflavi
18	191	10.7	224	1 G64383	riboflavin-specifi
19	184	10.3	216	1 C69129	riboflavin-specifi
20	180	10.1	224	1 G70339	riboflavin-specifi
21	163	9.1	213	1 S74011	probable riboflavi
22	146.5	8.2	178	1 F65033	hypothetical 20.0
23	142.5	8.0	151	2 G70377	conserved hypothet
24	142	7.9	148	2 D81879	probable cytosine
25	142	7.9	173	2 C64161	hypothetical prote
26	140	7.8	344	2 A64708	riboflavin biosynt
27	140	7.8	344	2 G71813	probable riboflavi
28	139	7.8	164	1 S74803	hypothetical prote
29	139	7.8	239	2 H81141	cytidine and deoxy

conserved hypothet
probable deoxycyti
hypothetical prote
dCMP deaminase (EC
dCMP deaminase (EC
riboflavin biosynt
hypothetical prote
probable cell cycl
hypothetical prote
cytosine deaminase
probable cytosine
dCMP deaminase hom
RIB2 protein - yea
hypothetical prote
probable RNA polym

ALIGNMENTS

RESULT 1

PN0100
riboflavin-specific deaminase ribG - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 04-Dec-1992 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
C:Accession: S45543; PN0100; E69692
R:Sorokin, A.; Zumstein, E.; Azevedo, V.; Ehrlich, S.D.; Serror, P.
submitted to the EMBL Data Library, November 1993
A:Reference number: S45533
A:Accession: S45543
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-361 <SOR>

A:CROSS-references: EMBL:L09228; NID:g410114; PIDN:AAA67481.1; PID:g410125
R:Mironov, V.N.; Perumov, D.A.; Krayev, A.S.; Stepanov, A.I.; Skryabin, K.G.
Mol. Biol. (Mosk.) 24, 256-261, 1990
A:Title: Unusual structure of Bacillus subtilis rib-operon regulatory region.
A:Reference number: PN0100; MUID:90271920
A:Accession: PN0100
A:Molecule type: DNA
A:Residues: 1-12 <MIR>
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogilwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Toso, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
A:Reference number: A69580; MUID:98044033
A:Accession: E69692
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-361 <KUN>
A:CROSS-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CAB14260.1; PID:g26347

C:Superfamily: Chlamydomonadales riboflavin-specific deaminase

Query Match 30.6%; Score 547; DB 2; Length 361;
Best Local Similarity 36.1%; Pred. No. 3.8e-34;
Matches 130; Conservative 68; Mismatches 134; Indels 28; Gaps 8;

C:Genetics:

A:Gene: ribD; ribG

C:Function:

A:Description: converts 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone 5'-phosphate into

A:Pathway: riboflavin biosynthesis

C:Superfamily: Chlamydophila pneumoniae riboflavin-specific deaminase

C:Keywords: hydrolase; riboflavin biosynthesis; zinc

F:55,80,89/Binding site: zinc (His, Cys, Cys) #status predicted

Query Match 27.7%; Score 495.5; DB 2; Length 372;

Best Local Similarity 34.3%; Pred. No. 3.2e-30;

Matches 124; Conservative 67; Mismatches 148; Indels 23; Gaps 9;

Qy 1 MDYAIQLANMVQGTGNNPPGVAVVNGRIVTGIGAHLRKGDKHAHQALDMAQNAEGA 60

Db 12 MORALDLAAKGOYTTPNPSVGCVLKNGEIVGEGFHFKAGOPHAERVALAQAGENAKGA 71

Qy 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKVYATKDNSLDTHGD--ETLRAHGIE--VEC 116

Db 72 TAYITLPCAHYGRTPPCALGLIEAGVVKVIAAMQDPNPQVAGRLKMLSDAGIESTVNL 131

Qy 117 VDDERASOLYODFFKAKAKOLPOITVKVSASLDGKQANDNGOSQWITNKVKQDYVKLRH 176

Db 132 LNDQ-AEKINGFLKMRQGMFFVOLKIAMSLDORTAMAGESKWTGPDARSVDQKMR 190

Qy 177 RHDVLTGRRTVELDDPOYTTRIQDG-----KNPIKIVILSKSGNIHFNOQIY 223

Db 191 KSSALLSTSTVIADDPNSLVNRWDEFFENLKTKEYKKEWLKRPVVRILDSQHRIOPTHKLF 250

Qy 224 QDESTPIWIYENP-NLTSNTHIE-IYILKSCDLTTLILHNLRYRGVGTLLVEAGPTTT- 280

Db 251 LTHS-PVWLVSSEPRDLTFPFCFQIIFPKENLLKELMRELGRQINTLWVEAGANLSG 309

Qy 281 SEFSYIYIDELIYAPKLIGSGNYQFYQNDVIEIPDANQFEIVHSELNQNKKTLR 340

Db 310 SLIDAKLVDELIIYIAPKLLGDNAR-GLCQLPNITKADAPLMQLEUIGQDDIKLTYT 368

Qy 341 KK 342

Db 369 PK 370

RESULT 5

riboflavin-specific deaminase NMB1817 [imported] - Neisseria meningitidis (group B strain

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000

C:Accession: B81039

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.;

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: B81039

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-369 <TET>

A:Cross-references: GB:AE002531; GB:AE002098; NID:g7227065; PIDN:AAF42152.1; PID:g722707

A:Experimental source: serogroup B, strain NMD58

C:Genetics:

A:Gene: NMB1817

C:Superfamily: Chlamydophila pneumoniae riboflavin-specific deaminase

Query Match 27.1%; Score 484; DB 2; Length 369;

Best Local Similarity 33.6%; Pred. No. 2.3e-29;

Matches 120; Conservative 68; Mismatches 151; Indels 18; Gaps 9;

Qy 1 MDYAIQLANMVQGTGNNPPGVAVVNGRIVTGIGAHLRKGDKHAHQALDMAQNAEGA 60

Db 12 MORALDLAAKGOYTTPNPSVGCVLKNGEIVGEGFHFKAGOPHAERVALAQAGENAKGA 71

Qy 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKVYATKDNSLDTHGD--ETLRAHGIE--VEC 116

Db 72 TAYITLPCAHYGRTPPCALGLIEAGVVKVIAAMQDPNPQVAGRLKMLSDAGIESTVNL 131

Qy 117 VDDERASOLYODFFKAKAKOLPOITVKVSASLDGKQANDNGOSQWITNKVKQDYVKLRH 176

Db 132 LNDQ-AEKINGFLKMRQGMFFVOLKIAMSLDORTAMAGESKWTGPDARSVDQKMR 190

Qy 177 RHDVLTGRRTVELDDPOYTTRIQDG-----KNPIKIVILSKSGNIHFNOQIY 223

Db 191 KSSALLSTSTVIADDPNSLVNRWDEFFENLKTKEYKKEWLKRPVVRILDSQHRIOPTHKLF 250

Qy 224 QDESTPIWIYENP-NLTSNTHIE-IYILKSCDLTTLILHNLRYRGVGTLLVEAGPTTT- 280

Db 251 LTHS-PVWLVSSEPRDLTFPFCFQIIFPKENLLKELMRELGRQINTLWVEAGANLSG 309

Qy 281 SEFSYIYIDELIYAPKLIGSGNYQFYQNDVIEIPDANQFEIVHSELNQNKKTLR 340

Db 310 SLIDAKLVDELIIYIAPKLLGDNAR-GLCQLPNITKADAPLMQLEUIGQDDIKLTYT 368

Qy 341 KK 342

Db 369 PK 370

Db 10 MENALRLAALGRFSTSPNPRVGCIVIAHSGQIVGQFHVKAGEPHAHVLRQAGEMAQGA 69

Qy 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKVYATKDNS--LDTHGDETFLRAHGEVEEC-V 117

Db 70 TAFVTLPCSHYGRTPPCAEALVRAGVSRVVAAMRDPNPLVAGKGLALLEAAGIKTECGL 129

Qy 118 DDERASOLYODFFKAKAKOLPOITVKVSASLDGKQANDNGOSQWITNKVKQDYVKLRH 177

Db 130 LEHQARELNKRGFLSRIERRRPFVRLKCAVSLDGKLTALSDGSSFWITGEDARADVQVLR 189

Qy 178 HDVLTGRRTVELDDPOYTTR-IOGKNPKIVILSKSGNIHFNOQIYODESTPIWIYEN 236

Db 190 SCAVLTGIGIVLADNPRNLNVRAPFTLRQPARIVLDSRLRPLDLSHLITDQOSPYYIATLE 249

Qy 237 PNLT-----SQNTHIEIY-LKSCDLTTLILHNLK-----RCVGTLLVEAGPTTTSEFSIY 286

Db 250 RNEDRLPYREHAHVIRILMPSETADSKIDLHLMLRLADEGFGIEMVEAGSELTSAF 309

Qy 287 YI-DEFILYAPKLI-GSGNYQFYQNDVIEIPDANQFEIVHSELNQNKKTLR 341

Db 310 NLADEIVLYRSPKILGSGKDLFSLLENRAALSAPPL--WTPVSSIEILGHDIKTVFRK 364

RESULT 6

D81984

probable diaminohydroxyphosphoribosylaminopyrimidine deaminase/phosphoribosylamino)ur

C:Species: Neisseria meningitidis

C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000

C:Accession: D81984

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Reference number: A81775; MUID:20222556

A:Accession: D81984

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-369 <PAR>

A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83933.1; PID:g737

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: ribD; NMA0644

C:keywords: oxidoreductase

Query Match 26.8%; Score 480; DB 2; Length 369;

Best Local Similarity 33.9%; Pred. No. 4.7e-29;

Matches 121; Conservative 69; Mismatches 149; Indels 18; Gaps 9;

Qy 1 MDYAIQLANMVQGTGNNPPGVAVVNGRIVTGIGAHLRKGDKHAHQALDMAQNAEGA 60

Db 10 MENALRLAALGRFSTSPNPRVGCIVIAHSGQIVGQFHVKAGEPHAHVLRQAGEMAQGA 69

Qy 61 TIYITLPCSHFGSTPPCVNKIIDCKIAKVYATKDNS--LDTHGDETFLRAHGEVEEC-V 117

Db 70 TAFVTLPCSHYGRTPPCAEALLRSGVTRVVAAMRDPNPLVAGKGLALLEAAGIKTECGL 129

Qy 118 DDERASOLYODFFKAKAKOLPOITVKVSASLDGKQANDNGOSQWITNKVKQDYVKLRH 177

Db 130 LEHQARELNKRGFLSRIERRRPFVRLKCAVSLDGKLTALSDGSSFWITGEDARADVQVLR 189

Qy 178 HDVLTGRRTVELDDPOYTTR-IOGKNPKIVILSKSGNIHFNOQIYODESTPIWIYEN 234

Db 190 SCAVLTGIGIVLADNPRNLNVRAPFTLRQPARIVLDSRLRPLDLSHLITDQOSPYYIATLE 249

Qy 235 -ENPNLTSNTHIEIY-LKSCDLTTLILHNLK-----RCVGTLLVEAGPTTTSEFSIY 286

Db 250 RDEKLRPYREHAHVIRILMPSETADGKIDLHLMLRLADEGFGIEMVEAGSELTSAF 309

Qy 287 YI-DEFILYAPKLI-GSGNYQFYQNDVIEIPDANQFEIVHSELNQNKKTLR 341

Db 310 NLADEIVLYRSPKILGSGKDLFSLLENRAALSAPPL--WTPVSSIEILGHDIKTVFRK 364

```

RESULT 7
S26201
riboflavin-specific deaminase (EC 3.5.4.-) - Escherichia coli
C:Species: Escherichia coli
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Apr-2000
C:Accession: S26201; F64770; S19178
R:Taura, T.; Ueguchi, C.; Shiba, K.; Ito, K.
Mol. Gen. Genet. 234, 429-432, 1992
A:Title: Insertional disruption of the nusB (ssyB) gene leads to cold-sensitive growth of Escherichia coli K-12.
A:Reference number: S26200; MUID:93024316
A:Accession: S26201
A:Molecule type: DNA
A:Residues: 1-367 <TAU>
A:Cross-references: EMBL:X64395; NID:g42147; PIDN:CAAA45735.1; PID:g581147
A:Experimental source: strain K-12
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, D.; Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: F64770
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-367 <BLAT>
A:Cross-references: GB:AE000148; GB:U00096; NID:gl786614; PIDN:AAC73517.1; PID:gl786616;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: ribD; ribG
A:Start codon: GTG
C:Function:
A:Description: converts 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone 5'-phosphate into 2,5-diamino-6-aminouracil
A:Pathway: riboflavin biosynthesis
C:Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase
C:Keywords: hydrolase; riboflavin biosynthesis; zinc
F:50,75,84/Binding site: zinc (His, Cys, Cys) #status predicted

Query Match 26.3%; Score 470; DB 2; Length 367;
Best Local Similarity 33.1%; Pred. No. 2.7e-28;
Matches 120; Conservative 58; Mismatches 135; Indels 50; Gaps 8;

QY 1 MDYAIQLANNVQGTGPNPVGAVVYVNEGRIIVGIGAHLRKGDHAEVQALDMAQQNAEGA 60
Db 7 MARALKLAQRGRFTTHPNPVGCVIVKDGIEVGEYHORAGEPHAEVHALRMAGEKAKGA 66

QY 61 TIVITLPCSHFGSTPPCVNKKIIDCKIAKVYA--TKDMSLDTHGDETLEHRAHGVIEC-V 117
Db 67 TAVITLPCSHHGRTPCCDALLAAGVARVVASMQDPNFQVAGRGYLRQQAGIDVSHGL 126

QY 113 EVECVDDEASQLYQDFFKAKAKQLPQITVKVSASLDGKQANDNGSQWITNKVEKQDVY 172
Db 127 MM-----SEAQLNKGFLKMRGTGFPYQLKGLASLDGRMTAMASGESQWITSPQARRDQ 181

QY 173 KLRHRDVAUTGRTVELDDPQVTR-----IQDGKPKIVLKSNGIHFN 219
Db 182 LLRAQSHAILTSATLADDPALTVRWSELDEQTAQLYPOQLRQPIRIVIDSQNRVTPV 241

QY 220 QQIYQDESTPIWLYTENP-----LTSNTHIEIILKSCDLTILHNLKRGV 268
Db 242 HRVQQPGTWFATQDSDSEWPEVTRTLPIPEKHGL-----DLVVMQLGKQKI 293

QY 269 GTLLVAPAGPTTSE-PSIYYIDFILIYAPKLGSGNYYQFYQTNVDVIEIPDANQF---E 324
Db 294 NSIWEAGPTLACALLAQGLVDELIVYIAPKLGSDAR-GLCTLPGLKLAADAPQFKKE 352

QY 325 IVH 327
Db 353 IRH 355

RESULT 8
G72026
riboflavin-specific deaminase CP0998 [imported] - Chlamydomophila pneumoniae (strains C
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: G72026; H81514
R:Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: G72026
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <ARN>
A:Cross-references: GB:AE001667; GB:AE001363; NID:g4377171; PIDN:AAD19009.1; PID:g437
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: H81514
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <REA>
A:Cross-references: GB:AE002257; GB:AE002161; NID:g7189902; PIDN:AAF38776.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: ribD; CP0998
C:Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase
```

RESULT 9

```

F70901
probable riboflavin-specific deaminase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70901
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: F70901
```

Db 72 ARGATAYVTLEPCSHDGRTPPCADALIAAGVARVVVVAAGDPNPQVNGRGLKLEKRAAGIEV 131

QY 115 ECVDDERASOLYQDFKRA-KAKOLPQITVKVSASLDCKQANDNGOSQWITNKEVKQDVK 173

Db 132 ATGVLEAAVROQAGFRSLVTRGRPHVYKYAMTLDCKVAALNEGNGPVSQPEARVMA 191

QY 174 LRRHDAVLTCRRITVELDDPYTTR-IDQGNPKPIKVLKSGNIIHFNOQ-----IY 223

Db 192 WRNEVDAAVAGTALLDNPNVRLGGDRPRAVLDPDEGHLPASARAVREGTVLVLR 251

QY 224 QDESTPTWITENPNLTSNQTHTIEIYLKSCDLTTLIHLNLYKRGVGLLVEAGPTTSE- 282

Db 252 EGRSTPL---ERDPRVTLVLAH-----SLOGALEQLAGLGVATLLLEGGPTLASAF 299

QY 283 FSIYVIDEFLYAPKLIG 301

Db 300 FEAGLIDELRVFAPKLIG 318

RESULT 11

S74377

riboflavin biosynthesis protein ribG - Synechocystis sp. (strain PCC 6803)

N;Alternate names: hypothetical protein slr0066

C;Species: Synechocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C;Accession: S74377

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis.

A;Reference number: S74322; MUID:97061201

A;Accession: S74377

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-368 <KAN>

A;Cross-references: EMBL:D64001; GB:AB001339; NID:g1001102; PIDN:BAA10295.1; PID:g1001102; The nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Genetics:

A;Gene: ribG

C;Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase

Query Match 23.1%; Score 413.5; DB 2; Length 368;

Best Local Similarity 32.3%; Pred. No. 5.3e-24;

Matches 108; Conservative 54; Mismatches 157; Indels 15; Gaps 5;

QY 1 MDYAIOIANVVGQGTGVNPPVGA VVYVNGRIGVIGAHRLKGRDKHAEVQALDMAQQNAEGA 60

Db 8 MRRLTLAKTAIGKATAPNPLVGSVIVQDEIVGGFHPQAGPHGEIFALWEAGDRAKGA 67

QY 61 TTYITLPECSHFSTPCPVNKIIDCKIAKVYATKDNS--LDTHGDETLEAHGIEVECDV 118

Db 68 TLVYNLEPCNHQGRTPCTEAIQAGIAKVVGVGMDPNPLVAGKISRLRQAGIEVKGV 127

QY 119 DERASOLYQDFERAKAK-QLPQITVKVSASLDCKQANDNGOSQWITNKEVKQDVKLHR 177

Db 128 EEACQRLNEAFRIKHQRPFGIFKVTAMTLDGKIATAQAHSSWVTSSSARHWVHQLRSQ 187

QY 178 HDVLTGRRITVELDDPYTTRIQDGKNPKIKVLKSGNIIHFNOQIYODESTPTWITENP 237

Db 188 QCAVILGNTVRRDNPLLTNHHGVEVNPPLVLSRLSLDLPLEAQLMDDVAKTLVITEKT 247

QY 238 NLTSNQTHTIEIYLKSCDLTTLIHLNLYKRGVGLLVEAGPTTSE-FSIYVIDE 291

Db 248 CDRNTLSHLEKLETVLVBQLTPLAVNEELYQRNCLQVLWECCGILAAEAIAAMGTQVK 307

QY 292 ILYYAPKLIGG-----SGNYQVQTNVDVIEPD 319

Db 308 HAFILAPKIIGVAAPTVPVGBELGFGQMTQALNLT 341

RESULT 12	
D81742	
Riboflavin-specific deaminase TC0103 [imported] - Chlamydia muridarum (strain Nigg)	
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn	
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000	
C:Accession: D81742	
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000	
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.	
A:Reference number: A81500; MUID:20150255	
A:Accession: D81742	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-396 <FET>	
A:Cross-references: GB:AE002277; GB:AE002160; NID:g7190127; PIDN:AAF38983.1; PID:g719013	
A:Experimental source: Strain Nigg (MoPn)	
C:Genetics:	
A:Gene: TC0103	
Query Match 23.0%; Score 411.5; DB 2; Length 396;	
Best Local Similarity 30.9%; Pred. No. 8.3e-24;	
Matches 104; Conservative 67; Mismatches 143; Indels 23; Gaps 8;	
QY 18	NPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQQNAEGATYITILEPCSHFGSTPP 77
DB 54	NPWVCVIVKNGCIIIGRWKHGIGSPHAEVCAFDQDTSSLVAGADVIVVTLPEPCCHFGRTTP 113
QY 78	CVNKIIDCKIAKVYA--TKDNLSDTHGDETLRAHGIIEVC-VDDERASQLYQDFPKA 134
DB 114	CVDLLIKSVSSVYIALLDPPRVCKRGVARLKEAGISVYVVGIGHEAKASLQPYLHQRE 173
QY 135	KQLPQITVKSASLDGKQANDNCSOWITNKEVKQDVYKLRHRHDAVLGRRRTVELDDPQ 194
DB 174	TGLPWWVMTAAASLDGQTSRRGISQISWISGEQARLDVGLRAESQAVIVGSRIVCLDNR 233
QY 195	YTTRIODG----KNPIKVLKSGNTHFNQOIYODESTPIWIYTNPNLTSTNQTHIE 247
DB 234	LSARMPGDLIEROPLRVVVDNRSGVPLDARVNPDSGNVLLATTE---QCSKEHIQKLE 290
QY 248	---IYLKS----CDLTTLIHLNLYKRGVTLVLAEGPTTTFSEF-SIYYIDEFLIYAPKL 299
DB 291	DRGVEVWKSPPQDVKRLQLYLAERGCQLVLEGGARLHSAFWERHLYNAGVIYNGPKF 350
QY 300	IGSGSNVQFYQTNVDVTEIPDANQFEIVHSELNQNPK 336
DB 351	LGDQGSFMLRLDQ--LCLDNAEHVKITKTFVLGDSVK 385
RESULT 13	
E71479	
probable riboflavin deaminase - Chlamydia trachomatis (serotype D, strain UW3/Cx)	
C:Species: Chlamydia trachomatis	
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 20-Apr-2000	
C:Accession: E71479	
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998	
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis	
A:Reference number: A71570; MUID:99000809	
A:Accession: E71479	
A:Status: preliminary	
A:Molecule type: DNA	
A:Residues: 1-375 <ARN>	
A:Cross-references: GB:AE001343; GB:AE001273; NID:g3329176; PIDN:AAC68325.1; PID:g332918	
A:Experimental source: serotype D, strain UW-3/Cx	
C:Genetics:	
A:Gene: ribD	
C:Superfamily: Chlamydia pneumoniae riboflavin-specific deaminase	
Query Match 22.8%; Score 408.5; DB 2; Length 375;	
Best Local Similarity 31.8%; Pred. No. 1.3e-23;	

Matches 108; Conservative 58; Mismatches 145; Indels 29; Gaps 9;	
QY 18	NPVGVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQQNAEGATYITILEPCSHFGSTPP 77
DB 29	NPWVCVIVKNGCIIIGRWKHGIGSPHAEVCAVDQDKCSLEGAIEFVTLPEPCCHFGRTTP 88
QY 78	CVNKIIDCKIAKVYA--TKDNLSDTHGDETLRAHGIIEVC-VDDERASQLYQDFPKA 134
DB 89	CVDLLIKSVAAVYVGLLDDPPRVCKRGVARLQAAGIPVYVGVSGSEAKTSLQPYLYQRE 148
QY 135	KQLPQITVKSASLDGKQANDNCSOWITNKEVKQDVYKLRHRHDAVLGRRRTVELDDPQ 194
DB 149	RGLPWWVMTAAASLDGQTSRRGISQISWISGEQARLDVGLRAESQAVIVGSRIVCLDNR 208
QY 195	YTTRIODG----KNPIKVLKSGNTHFNQOIYODESTPIWIYTNPNLTSTNQTHIE 250
DB 209	LSARMPGDLIEROPLRVVVDNRSGVPLDARVNPDSGNVLLATTE---TTOCPKEYIQ 262
QY 251	-LK-----SCDLTTLIHLNLYKRGVTLVLAEGPTTTFSEF-SIYYIDEFLIY 296
DB 263	KLKDLGVEVWESSSHQVDLKLRLYLAERGCQLVLEGGARLHSAFWERHLYNAGVIYNG 322
QY 297	PKLIGSGSNVQFYQTNVDVTEIPDANQFEIVHSELNQNPK 336
DB 323	PKPLGDSQFMLRLDQ--LSLVTAEHVRITETSLVRDSVK 360
RESULT 14	
T12994	
Riboflavin biosynthesis protein ribG homolog T21L8.140 - Arabidopsis thaliana	
C:Species: Arabidopsis thaliana (mouse-ear cress)	
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 26-May-2000	
C:Accession: T12994	
R:Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; submitted to the Protein Sequence database, July 1999	
A:Reference number: Z17586	
A:Accession: T12994	
A:Molecule type: DNA	
A:Residues: 1-599 <CHO>	
A:Cross-references: EMBL:AL096860; GSPDB:GN00061; ATSP:T21L8.140	
A:Experimental source: cultivar Columbia; BAC clone T21L8	
C:Genetics:	
A:Gene: ATSP:T21L8.140	
A:Map position: 3	
A:Introns: 119/3; 164/3; 198/1; 240/1; 344/3; 384/3; 437/1; 513/3	
Query Match 17.9%; Score 320; DB 2; Length 599;	
Best Local Similarity 25.9%; Pred. No. 1.3e-16;	
Matches 91; Conservative 69; Mismatches 145; Indels 46; Gaps 8;	
QY 4	ATQLANWVGQTVNPPVCAV-VNIEGRIVGICAHLRKGDHAEVQALDMAQQNAEGATI 62
DB 39	ADLSEMSAGLTSPPHFCVTSATSSGKVAGEGYLYAQGTKPAELAVAAAGEFSRGATA 98
QY 63	YITLPCSHFGSTPPCVNKKIIDCKIAKVYATKNSLDTHGD--ETLRAHGEVECVDE 120
DB 99	YLNMEPGDCHGD-HTAVSALVQAGIERVVVGIHPLQLHLSAIRELSHGIEVNVLGED 157
QY 121	RASQLYQDFPKA-----KAKQLPQITVKSASLDGKQANDNCSOWITNKEVKQ 169
DB 158	FESKLEDAKSCLLYNAPLIHRACSRVPFVLKYAMTLDGKIAASGSHAAMISSKLSRT 217
QY 170	DYVKLRHRHDAVLGRRRTVELDDPQYTTRIODGKNPIKVLKSGNTHFNQOIYODESTP 229
DB 218	RVFELRGSDAIVGVGTVRQDDPLRTARHGQHTPTTRIVMTQSLDL-----PEKAN 269
QY 230	IWIYTNPNLTSTNQ-----THIEIYLKSCDLTTLIHLNLYKRGVGTLLVEAG 276
DB 270	LNDVSESVITVVTQCGARKSFQKLLASKGVEVVEFDMLNPREVMEYFHLRGVLSILWECG 329
QY 277	PT-TTSEFSIYIDEFLIYAPKLIGGS-----GNVQFYQTNVDVTEI 317

Db 330 GTLAASAISSSVIHKVVAFVAPKIIIGSKAPSPVGD LGMVEMTQALNLIDV 360

RESULT 15

T10638

hypothetical protein T13K14.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Apr-2000

C:Accession: T10638

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke

submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16991

A:Accession: T10638

A:Molecule type: DNA

A:Residues: 1-363 <BEV>

A:Cross-references: EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.120

A:Experimental source: cultivar Columbia; BAC clone T13K14

C:Genetics:

A:Gene: ATSP:T13K14.120

A:Map position: 4

A:Introns: 60/3; 160/2

C:Superfamily: Chlamydomonada pneumoniae riboflavin-specific deaminase

Query Match 16.8%; Score 300; DB 2; Length 363;
Best Local Similarity 28.4%; Pred. No. 2.1e-15;
Matches 83; Conservative 61; Mismatches 118; Indels 30; Gaps 8;

QY 1 MDYAIQLANWVGOTGVNPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQOAE 60
Db 15 MRKVELAKRAIGCTSPNPMWGCIVVKGDI VGGFHPKAGQPHAEVFAIRDAGELAE 74
QY 61 TIYTILEPCSHFGSTPCVANKIIDCKIAKVYATKDNS--LDTHGDETLRAHGIEVE-CV 117
Db 75 TAVVSLPCNHYGRTPTCTEALIKAKVRRVIGMVDNPVIFSSGISRLKADGIDVTVS 134
QY 118 DDERASQLYODFFKAKAKQLPQITVKVSASLD-----CKQANDNGCSQWITNKEVKQD 170
Db 135 EEELCKRMNEGFTHRMLTGTGPFALRYSMVSGCLLDKICGGASDSG-----G 182
QY 171 VY-KLRHRHDA-VLTGRTVELDDPQVTTTIOCKNPKIVL---SKSGNTHFNQIYQD 225
Db 183 YSKLLQYDAIILSSLSDELSSISQEAINVSIQPIQIIVASNAQOSHILASSHTVEE 242
QY 226 ESTPIWIYTNPNLTS---NOTHIEIILYKSCDLTTILHNLKRGVGTLLVE 274
Db 243 SGPKVVVFTAKESVAESGSSGVETVVKLEKINLDSILDYCYNRLGCSVLDD 294

Search completed: November 14, 2000, 10:08:13
Job time: 91 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:07:31 ; Search time 16.17 Seconds
(without alignments)
675.708 Million cell updates/sec

Title: US-08-978-456-2

Perfect score: 1789

Sequence: 1 MDYATQLANMVGQGTGVNPP.....FEIVHSELLNQVKLTLRKK 342

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	547	30.6	361	1 RIBD_BACSU	P17618 b riboflavi
2	545	30.5	356	1 RIBD_AQUAE	O66534 a riboflavi
3	513	28.7	371	1 RIBD_BACAM	P70814 b riboflavi
4	495.5	27.7	372	1 RIBD_HAEIN	P44326 h riboflavi
5	470	26.3	367	1 RIBD_ECOLI	P25539 e riboflavi
6	435	24.3	376	1 RIBD_CHLPN	O92735 c riboflavi
7	434.5	24.3	339	1 RIBD_MXTU	P71677 m riboflavi
8	413.5	23.1	368	1 RIBD_SYNY3	O55158 s riboflavi
9	393	22.0	376	1 RIBD_ACTPL	P50853 a riboflavi
10	348	19.5	337	1 RIBD_CORAM	O24750 c riboflavi
11	195	10.9	219	1 RIB7_ARCFU	O28272 archaeoglob
12	191	10.7	224	1 RIB7_METJA	O58085 methanococc
13	184	10.3	216	1 RIB7_METH	O26337 methanobact
14	163	9.1	213	1 RIB7_SULSO	P95872 sulfobolob
15	146.5	8.2	178	1 YFHC_ECOLI	P30134 escherichia
16	142.5	8.0	151	1 Y903_AQUAE	O67050 a quifex aeo
17	142	7.9	173	1 YFHC_HAEIN	P44931 haemophilus
18	135	7.5	161	1 YAAJ_BACSU	P21335 bacillus su
19	128.5	7.2	158	1 FCYL_YEAST	O12178 saccharomyc
20	123.5	6.9	193	1 DCTD_BPT4	P16006 bacterioph
21	121.5	6.8	188	1 DCTD_BPT2	P00814 bacterioph
22	117	6.5	244	1 RIB7_YEAST	P33312 saccharomyc
23	104	5.8	591	1 RIB2_YEAST	O12362 saccharomyc
24	101	5.6	147	1 YLXG_VIBFI	P33968 vibrio fisc
25	100.5	5.6	150	1 FCAL_CANAL	P78594 candida alb
26	100.5	5.6	611	1 RPST_TREPA	O83506 treponema p
27	100	5.6	792	1 SYL_MYCGE	P7508 mycoplasma
28	97	5.4	148	1 Y831_RICPR	O92cc6 rickettsia
29	97	5.4	3343	1 YOG7_CAEEL	P34616 caenorhabdi
30	95.5	5.3	4644	1 DYHC_RAT	P38650 rattus norv
31	95	5.3	770	1 A4_MOUSE	P12023 mus musculu
32	94	5.3	405	1 EFTU_DEISP	P33168 deinonema s
33	94	5.3	559	1 FKB7_WHEAT	O43207 triticum ae

RESULT 1

ID	RIBD_BACSU	STANDARD;	PRT;	361 AA.
AC	P17618;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:			
DE	DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)			
DE	(RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-			
DE	PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].			
GN	RIBD OR RIBG.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / MABBURG;			
RX	MEDLINE; 95020538.			
RA	Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;			
RT	"The organization of the Bacillus subtilis 168 chromosome region			
RT	between the spoVA and serA genetic loci, based on sequence data.";			
RL	Mol. Microbiol. 10:385-395(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / SHGW;			
RA	Mironov V.N.;			
RL	Thesis (1989), USSR Academy of Sciences, Russia.			
RN	[3]			
RP	SEQUENCE OF 1-12 FROM N.A.			
RC	STRAIN=168 / SHGW;			
RA	Mironov V.N., Perumov D.A., Kraev A.S., Stepanov A.I., Skryabin K.G.;			
RT	"Unusual structure of the regulatory region of the riboflavin			
RT	biosynthesis operon in Bacillus subtilis.";			
RL	Mol. Biol. (Mosk) 24:256-261(1990).			
RN	[4]			
RP	CHARACTERIZATION.			
RC	MEDLINE; 97221604.			
RA	Richter G., Fischer M., Krieger C., Eberhardt S., Luttmann H.,			
RA	Gerstenschlager I., Bacher A.;			
RT	"Biosynthesis of riboflavin: characterization of the bifunctional			
RT	deaminase-reductase of Escherichia coli and Bacillus subtilis.";			
RL	J. Bacteriol. 179:2022-2028(1997).			
CC	-1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-			
CC	5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-			
CC	PYRIMIDINEDIONE 5'-PHOSPHATE.			
CC	-1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-			
CC	PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-			
CC	PHOSPHORIBOSYLAMINO)URACIL + NH(3).			
CC	-1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +			
CC	NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.			
CC	-1- COFACTOR: ZINC (BY SIMILARITY).			
CC	-1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.			
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND			
CC	DEOXYCYTIDYLATE DEAMINASES FAMILY.			

34	93.5	5.2	321	1	PE38_NPVAC	P23801 autographa
35	93	5.2	770	1	A4_HUMAN	P05067 homo sapien
36	93	5.2	4196	1	DYHC_SCHPO	O13290 schizosacch
37	92.5	5.2	419	1	CARV_CANAL	P10977 candida alb
38	92.5	5.2	4639	1	DYHC_DROME	P37276 drosophila
39	92	5.1	727	1	EF2_METVA	P09604 methanococc
40	92	5.1	738	1	PARC_RICPR	Q9ze79 rickettsia
41	92	5.1	770	1	A4_RAT	P08592 rattus norv
42	92	5.1	1233	1	YFI6_YEAST	Q43597 saccharomyc
43	91.5	5.1	751	1	A4_SAIISC	Q95241 saimiri sci
44	90.5	5.1	378	1	NIR_ACHCY	P25006 achromobact
45	90.5	5.1	982	1	P115_MYCGE	P47540 mycoplasma

ALIGNMENTS

```

CC CC      -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC CC      REDUCTASE FAMILY.
CC CC      -----
CC CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC      the European Bioinformatics Institute. There are no restrictions on its
CC CC      use by non-profit institutions as long as its content is in no way
CC CC      modified and this statement is not removed. Usage by and for commercial
CC CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC      or send an email to license@isb-sib.ch).
CC CC      -----
CC CC      EMBL; L09228; AAA67481.1; -.
CC CC      EMBL; X51510; CAA35878.1; -.
CC CC      EMBL; Z99116; CAB14260.1; -.
CC CC      PIR; A35711; A35711.
CC CC      PIR; S45543; S45543.
CC CC      SUBTILIST; BG10518; RIBD.
CC CC      INTERPRO; IPR002125; -.
CC CC      INTERPRO; IPR002734; -.
CC CC      PFAM; PF01872; RibD_C; 1.
CC CC      PFAM; PF00383; dCMP_cyt_deam; 1.
CC CC      PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
CC CC      Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
CC CC      Multifunctional enzyme.
CC CC      FT DOMAIN 1 144 DEAMINASE.
CC CC      FT DOMAIN 145 361 REDUCTASE.
CC CC      FT METAL 49 49 ZINC (BY SIMILARITY).
CC CC      FT METAL 74 74 ZINC (BY SIMILARITY).
CC CC      FT METAL 83 83 ZINC (BY SIMILARITY).
CC CC      SQ SEQUENCE 361 AA; 39305 MW; DA836930BFDECA3B CRC64;

Query Match 30.6%; Score 547; DB 1; Length 361;
Best Local Similarity 36.1%; Pred. No. 9.1e-35;
Matches 130; Conservative 68; Mismatches 134; Indels 28; Gaps 8;

QY 1 MDAIQLANNVQQTGVNPPVGAHVNVNNEGRVIGIGIAHLRKGDKHAEVQALDMAQQAEGA 60
DB 6 MKLALDLAQGEQTESNPLVGAVVVDQIVGMGAHLKYGAEHAEVHAHMAHAEGA 65

QY 61 TYITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETIRAHGIEV-ECV 117
DB 66 DIYVTLEPCSHYKTPPCAEILINSIGIKRVFVAMRPNPLVAGRGISMKEAGIEVREGI 125

QY 118 DDERASQYODFFKAKAKOLPOITVKVSASLDGKOANDGQOWINTKEVKQDVYKLRH 177
DB 126 LADQAEELNEKELHFHRTGLPYVTLKAAASLDGKIATSTGDSKWTSEARQDAQYRKT 185

QY 178 HDAVLTGRTVELDDPQYTRIQD-GKNPTKVLKSGNIHFNQIYODESTPIWIYT-- 234
DB 186 HQSILVGVGTVRADNPSLTCRLPNVTQKQVRVILDTVLSIPEDAKVICQIAPTWIFTTA 245

QY 235 -----ENPNLTSTNQTHIEIYLKSCDLTILHLNLYKRGVGTLLVEAGPTTSEF-SIYI 288
DB 246 RADEEKRLKLSAFGVNIFTLETRIQIPDVLIKALBEGMSVVEGGSVAHGSFVKEGCF 305

QY 289 DEFILYAPKLGG-----SG-NYQFYQTNDVIPDANOFIHSVSELLNQNVKLTLR 340
DB 306 QEIIFFAPKLGGTHAPSLISGEGFSQMKDVPFLQFTDITQ-----IGDKILTKAK 357

RESULT 2
RIBD.AQAE
ID RIBD.AQAE STANDARD; PRT; 356 AA.
AC 066534.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG OR AQ_138.

```

```

OS Aquifex aeolicus.
CC Bacteria; Aquificales; Aquificaceae; Aquifex.
CC [1]
RN SEQUENCE FROM N.A.
CC STRAIN-VF5;
CC MEDLINE; 98196666.
CC Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
CC Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
CC Feldman R.A., Short J.M., Olson G.J., Swanson R.V.:
CC "The complete genome of the hyperthermophilic bacterium Aquifex
CC aeolicus."
CC Nature 392:353-358(1998).
CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5'-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000675; AAC06487.1; -.
CC INTERPRO; IPR002125; -.
CC INTERPRO; IPR002734; -.
CC PFAM; PF01872; RibD_C; 1.
CC PFAM; PF00383; dCMP_cyt_deam; 1.
CC PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
CC Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
CC Multifunctional enzyme.
CC KW DOMAIN 1 148 DEAMINASE.
CC FT DOMAIN 149 356 REDUCTASE.
CC FT METAL 53 53 ZINC (BY SIMILARITY).
CC FT METAL 78 78 ZINC (BY SIMILARITY).
CC FT METAL 87 87 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 356 AA; 39628 MW; 7775515B1E4D02BD CRC64;

Query Match 30.5%; Score 545; DB 1; Length 356;
Best Local Similarity 40.5%; Pred. No. 1.3e-34;
Matches 125; Conservative 54; Mismatches 122; Indels 8; Gaps 5;

QY 1 MDAIQLANNVQQTGVNPPVGAHVNVNNEGRVIGIGIAHLRKGDKHAEVQALDMAQQAEGA 60
DB 10 MKLALSALKRKGTHPNPTVGAVVKEGKIVGLGHERAGRPHAEVMAVGQAGEKAKGA 69

QY 61 TYITILEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETIRAHGIEV-ECV 117
DB 70 TLYVTLEPCTHFGTRTPCCTDAIIRSGIKRVVAVATLDPNPLMSGKGVKELRNAGIEVDVG 129

QY 118 DDERASQYODFFKAKAKOLPOITVKVSASLDGKOANDGQOWINTKEVKQDVYKLRH 177
DB 130 CBEAEARELNEDEFTYITQERPIITLKAQTLGKALTLGSSKWTSEARQDAHILRE 189

QY 178 HDAVLTGRTVELDDPQYTRT-IQDGNPKIVLTKSGNIHFNQIYODESTPIWIYTEN 236
DB 190 ATAVLVGVNTVIKDDPHLTVRFVPTKEQVRIILDPLEVLPSAKVLNTEEAPTIVITKK 249

QY 237 PNLTSTNQTHIEIYLKSCDLTILHLNLYKRGVGTLLVEAGP-TTTSFSESIYIDEFTI 292
DB 237 PNLTSTNQTHIEIYLKSCDLTILHLNLYKRGVGTLLVEAGP-TTTSFSESIYIDEFTI 292

```

Db 250 ENERAEKLEKGVOLILKGFNLKILKLEIMHLMVEGGPRTLTSLFKEGFFDRIV 309

Qy 293 LYPAPKLIG 301

Db 310 VFIAPKMG 318

RESULT 3

ID RIBD_BACAM STANDARD; PRT; 371 AA.
AC P70814;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG.
OS Bacillus amyloliquefaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A 50.
RX MEDLINE; 97376523.
RA Gusarov I.I., Krenova R.A., Podcharniaev D.A., Iomantas I.U.V.,
RA Abalokina E.G., Stoinova N.V., Perumov D.A., Kozlov I.U.I.;
RT "Riboflavin biosynthetic genes in Bacillus amyloliquefaciens: primary
structure, organization and regulation of activity";
RL Mol. Biol. (Mosk) 31:446-453(1997).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X95955; CAA65189.1; -;
DR INTERPRO; IPR002125; -;
DR INTERPRO; IPR002734; -;
DR PRAM; PF01872; RibD_C.1.
DR PRAM; PF00383; dCMP_cyt_deam.1.
DR PROSITE; PS00903; CYT_dCMP_DEAMINASES; 1.
KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 144 DEAMINASE.
FT METAL 145 371 REDUCTASE.
FT METAL 49 49 ZINC (BY SIMILARITY).
FT METAL 74 74 ZINC (BY SIMILARITY).
FT METAL 83 83 ZINC (BY SIMILARITY).
SQ SEQUENCE 371 AA; 40166 MW; 1A90722056A6BBA9 CRC64;

Query Match 28.7%; Score 513; DB 1; Length 371;
Best Local Similarity 34.9%; Pred. No. 3.8e-32;
Matches 122; Conservative 62; Mismatches 154; Indels 12; Gaps 5;

Qy 1 MDVAIQANMVQGTGNPPVGVAVVNEGRIVGTGAHLRKGDKHAEOALDMAQONAEGA 60
Db 6 MNTAIELARRGEGTQPNPLVGVAVVVKRQIVGMGAHLQYGEAHEAHEAIAINMAGSLAKGA 65
Qy 61 TIYITLPCSHFGSTPPCVNKLIDCKIAKVYATKDNS--LDTHGDETLRAHGLEVEVC-V 117
Db 66 DLVITLPCSHFGSTPPCAELIMKSGIKRPIAIVEDNPLVAGKGTMLERAGLEVITGL 125
Qy 118 DDERASOLYQDFKAKAKAQPOITVKVVSASLDGKQANDNGOSOWITNKVKODVYKLRHR 177
Db 126 LRQAAEELNKMFLHFMRTGLPYVTLKKAASLDGKTATETGDSKWITSEARLDAQVRS 185
Qy 178 HDAVLTRGRVVELDDPYTTRIQQG-KNPIKVLKSGNTHFNQOIYODESTPIWITEN 236
Db 186 HQRLVAGTVKADNPISLTCRLPDVAKQPVRLDITKLTVPETANVLTDGAAPTWIFTAA 245
Qy 237 PNLTSNTHIEIIVLKSCDLTT-----ILHNLKRGVGTLLVEAGPTTISEF-SIYYI 288
Db 246 GSDVRKKDRLTALGKIVFTLETDRHIPEVLSILAENGIMSVYEGGASVHGSEVKGACF 305
Qy 289 DEFILYVAPKLIGGNGVQFYQTNVDVIEIPDANQFETVHSELLMONVKLT 338
Db 306 DELHFYFAPILIGTGLAPSLISGEGFQSMKDVPHLQFTQITQIGDPDKLT 355

RESULT 4

ID RIBD_HAEIN STANDARD; PRT; 372 AA.
AC P44326;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG OR HI0944.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20;
RX MEDLINE; 95350630.
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL +
CC NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; U32775; AAC22598.1; -;
 DR TIGR; HI0944; -;
 DR INTERPRO: IPR002125; -;
 DR INTERPRO: IPR002734; -;
 DR PFAM: PF01872; RibD_C; 1.
 DR PFAM: PF00383; dCMP_cyt_deam; 1.
 DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
 DR Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
 KW Multifunctional enzyme.
 FT DOMAIN 1 150 DEAMINASE.
 FT DOMAIN 151 372 REDUCTASE.
 FT METAL 55 55 ZINC (BY SIMILARITY).
 FT METAL 80 80 ZINC (BY SIMILARITY).
 FT METAL 89 89 ZINC (BY SIMILARITY).
 SQ SEQUENCE 372 AA; 40931 MW; C7504EDD8B86F7D CRC64;

Query Match 27.7%; Score 495.5; DB 1; Length 372;
 Best Local Similarity 34.3%; Pred. No. 8.2e-31;
 Matches 124; Conservative 67; Mismatches 148; Indels 23; Gaps 9;

QY 1 MDYAIQLANNVQGTGVNPPVGVAVVYNEGRIVGIGHLRKGDKHAQVQALDMAQNAEGA 60
 DB 12 MQRALDLAAKQGVYTPNPVSGCVLVKNGEIVGEGFHFKAGOPHAERVALAQAGENAKGA 71
 QY 61 TIVITLPCSHFGSTPPCKNIIDCKIAKVATKNSLDTHGD--ETLRAHGIE--VEC 116
 DB 72 TAVTLEPCAHYGRTPPCALGLIEAGVVKVIAAMQDPNPQVAGKGLKMLSDAGIESTVNL 131
 QY 117 VDDEASQLYQDFKAKAKQLPQITVKVSASLDGKQANDGOSQMTNKEVKQDVYKLRH 176
 DB 132 LNDQ-AEKINKGLKRMQGMPPVQLKLAMSLDGRMTAMAGESKWTGPDARSVDQKMR 190
 QY 177 RUDAVLTGRTEVLDPOYTRTDG-----KNPIKVLKSGNHFNQOY 223
 DB 191 KSSALLSTVTIADDPNSLVNRWDEFPENLKTEYKKEWLRQPVRVILDSOHRIQPTHKLF 250
 QY 224 QDSTPIWITENP-NLTSNQTHIE-IIVLKSCDLTLHLNLYKRGVGTLLVEAGPTT- 280
 DB 251 LTHS-PVWLVSSEPRDLTGPDCEQIIFPKENLLKELRELKRGKQINTLWVEAGANLSG 309
 QY 281 SEFSYIYIDEFILYAPKLIGSGNYQFYOTNDVIEIPDANOFIEVHSELLNONVKLTLR 340
 DB 310 SLIDAKLVDELIIYIAPKLIGDNAR-GLCQLPNLTKLADAPLWQLNELEIGDIDKLTYT 368
 QY 341 KK 342
 DB 369 PK 370

RESULT 5
 RIBD_ECOLI
 ID RIBD_ECOLI STANDARD; PRT; 367 AA.
 AC P25539;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
 DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
 DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
 DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
 GN RIBD OR RIBG.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RC [1]

RP SEQUENCE FROM N.A.
 RX STRAIN-K12;
 CC MEDLINE; 93024316.
 RA Taura T., Ueguchi C., Shiba K., Ito K.;
 RT "Insertional disruption of the nusB (ssyb) gene leads to cold-
 RT sensitive growth of Escherichia coli and suppression of the secY24
 RT mutation.";
 RL Mol. Gen. Genet. 234:429-432(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN-K12 / MG1655;
 CC MEDLINE; 97426617.
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
 RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
 RA Lew H., Lin D., Mamath A., Oefner P., Schramm S., Davis R.W.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE; 97221604.
 RA Richter G., Fischer M., Krieger C., Eberhardt S., Luttgen H.,
 RA Gerstenschlager I., Bacher A.;
 RT "Biosynthesis of riboflavin: characterization of the bifunctional
 RT deaminase-reductase of Escherichia coli and Bacillus subtilis.";
 RL J. Bacteriol. 179:2022-2028(1997).
 CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
 CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
 CC PYRIMIDINEDIONE 5'-PHOSPHATE.
 CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
 CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
 CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
 CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
 CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
 CC -!- COFACTOR: ZINC (BY SIMILARITY).
 CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
 CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
 CC REDUCTASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; X64395; CAA45735.1; -;
 DR EMBL; AE000148; AAC73517.1; -;
 DR EMBL; U82664; AAB40170.1; -;
 DR PIR; S26201; S26201.
 DR ECOGENE; EG11321; RIBD.
 DR INTERPRO; IPR002125; -;
 DR INTERPRO; IPR002734; -;
 DR PFAM: PF01872; RibD_C; 1.
 DR PFAM: PF00383; dCMP_cyt_deam; 1.
 DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
 KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
 KW Multifunctional enzyme.
 FT DOMAIN 1 145 DEAMINASE.
 FT DOMAIN 146 367 REDUCTASE.
 FT METAL 50 50 ZINC (BY SIMILARITY).
 FT METAL 75 75 ZINC (BY SIMILARITY).
 FT METAL 84 84 ZINC (BY SIMILARITY).
 SQ SEQUENCE 367 AA; 40338 MW; B19CEFA74D48D14D CRC64;

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.:
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PROSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PROSPHORIBOSYLAMINO)URACIL + NADPH.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; Z80108; CAB02188.1; -.
DR TUBERCULIST; RV1409; -.
DR INTERPRO; IPR002125; -.
DR INTERPRO; IPR002734; -.
DR PFAM; PF01872; RIBD_C; 1.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
DR Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 152 DEAMINASE.
FT METAL 57 339 REDUCTASE.
FT METAL 57 57 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
FT METAL 91 91 ZINC (BY SIMILARITY).
SQ SEQUENCE 339 AA; 35366 MW; 1DC0A1B8E7EC08B4 CRC64;

Query Match 24.3%; Score 434.5; DB 1; Length 339;
Best Local Similarity 35.0%; Pred. No. 3.3e-26;
Matches 109; Conservative 57; Mismatches 114; Indels 31; Gaps 9;
QY 1 MDYAIQLANNVQGTGVNPPVGVAVVNV-EGRIYIGAHLRKDKHAEVQALDMAQNAEG 59
DB 13 MGLAIEHSYQVGGTKTTPKPPGVIVDPNGRIYVAGGTEPAGGDHAEVVALRRAGSLAG 72
QY 60 ATIVITLEPCSHRGSPPCVKNKIIDCKIAKVYATKD-NSLDTHGDETLRHAGGIEVE--C 116
DB 73 AIVVWTFEPCNHGKTPPCVNLAIERVGTIVYAVADPNAGIAGGAGRLSAGLQVRSVG 132
QY 117 VDDERASQLYQDFFKAKAKOLPOITVKVSASLDGKANDNGSQWITNKVEVKDYKLRH 176
DB 133 LAEQVAAGPLREWLHKQRTGLPHVTWKYATSIDGRSAADGSSQWISSEARLDLHRRRA 192
QY 177 RHDVITGRTVELDDPQYTRTQDG-----KNPIKIVLSKSGNIHFNQIQDESTPIWI 232
DB 193 IADAILVGTGLVADDPALFARLADGSLAPQQLRVVWCKR-DIPPEARVLNDEARTMI 251
QY 233 YTEPNLUTSNOTHIEIYLKSCDLTILHNLKRGVGTLLVEAGPTTSTSEF-SIYVIDEF 291
DB 252 RTHEP-----MEVLRALS-DRTDV-----LLEGPTLAGAFLRAGAINRI 290

QY 292 ILYYAPKLIQG 302
DB 291 LAYVAPILLGG 301
RESULT 8
RIBD_SYNY3 STANDARD; PRT; 368 AA.
AC Q55158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR SLR0066.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96127529.
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.:
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -!- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -!- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PROSPHORIBOSYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PROSPHORIBOSYLAMINO)URACIL + NADPH.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.
CC -----
DR EMBL; D64001; BAA10295.1; -.
DR INTERPRO; IPR002125; -.
DR INTERPRO; IPR002734; -.
DR PFAM; PF01872; RIBD_C; 1.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
DR Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 146 DEAMINASE.
FT METAL 51 368 REDUCTASE.
FT METAL 51 51 ZINC (BY SIMILARITY).
FT METAL 76 76 ZINC (BY SIMILARITY).
FT METAL 85 85 ZINC (BY SIMILARITY).
SQ SEQUENCE 368 AA; 39995 MW; BCDABED0916B50BC CRC64;

Query Match 23.1%; Score 413.5; DB 1; Length 368;
Best Local Similarity 32.3%; Pred. No. 1.5e-24;
Matches 108; Conservative 54; Mismatches 157; Indels 15; Gaps 5;
QY 1 MDYAIQLANNVQGTGVNPPVGVAVVNVGRIYIGAHLRKDKHAEVQALDMAQNAEGA 60


```

RX MEDLINE: 98049343.
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.J., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -!- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF000964; AAB89247.1; -.
CC TIGR: AF2007; -.
CC DR INTERPRO: IPR002734; -.
CC DR PFAM: PF01872; RibD_C; 1.
CC DR Hypothetical protein: Riboflavin biosynthesis; Oxidoreductase; NADP.
KW SEQUENCE 219 AA; 24578 MW; 30DEA579024CF625 CRC64;
SQ -----
Query Match 10.9%; Score 195; DB 1; Length 219;
Best Local Similarity 29.8%; Pred. No. 3.7e-08;
Matches 54; Conservative 36; Mismatches 75; Indels 16; Gaps 4;
QY 138 PQITVKVVASLDGKQANDGSGOWITNKEVKQDVYKLRHRHDVLTGRRTVELDDPQYTT 197
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 PYVFNVAASLDGKISDESRRQLRISCEEDIRIVDRLRABSDAIVMGITGTVLADDPRLTV 62
QY 198 -----RTDCK--NPIKVLKSGNIHFNOOIYODESTPIWIYT-----ENPNLTSN 242
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 KSAELREKRODKKEPNLRPLRVVDSRCRPLTARILNDEARTLVAVSRAPKEEVKVK 122
QY 243 QTHIEIYILKSCDITTLNLNLYKRGVGTLLVEAGPTTTSF-FSIYYIDEFLYYAPKLIG 301
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 123 VAENVGGERVELSALLEFLHRKGVRLVVEGGTGLISSLSQNLDVEIRIYVGPFIIG 182
QY 302 G 302
DQ :
Db 183 G 183

RESULT 12
RIB7_METJA
ID RIB7_METJA STANDARD; PRT; 224 AA.
AC Q58085;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PUTATIVE 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE
DE (EC 1.1.1.193) (HTP REDUCTASE).
GN MJ0671.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;

```



```

RX MEDLINE; 96337999.
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervilange A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodex A.,
RA Scott J.L., Geobaghen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton H.-P., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67514; AAB98665.1; -.
CC TIGR; MJ0671; -.
CC INTERPRO: IPR002734; -.
CC PFAM; PF01872; RibD.C; 1.
CC KW Hypothetical protein; Riboflavin biosynthesis; Oxidoreductase; NADP.
CC SEQUENCE 234 AA; 25037 MW; 4D8C15CE291E89D8 CRC64;
CC
CC Query Match 10.7%; Score 191; DB 1; Length 224;
CC Best Local Similarity 28.0%; Pred. No. 7.8e-08;
CC Matches 61; Conservative 41; Mismatches 82; Indels 34; Gaps 9;
CC
CC Qy 138 POITVKVSASLDGKQANDNGOSQITNKEVKQDVYKLRHRHDAVLGTGRRTVELDDPQVTT 197
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 9 PYIISNMGWTLGDKLATINDSR-ISCEDLIRVHKIRANVDGIMVGIGTVLKDDPRLTV 67
CC
CC Qy 198 -RIQDQKNPIKVLKSGNIHFNQIYQDESTPIWIYTENPNLTSNQTHTIEI-----IYL 251
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 68 HKIKSDRNPRIVVDSKURVPLNARVLNKKDAITATTEDTN-EEKEKKIKILEDGMGEV 126
CC
CC Qy 252 KSC-----DLTILNLYKRGVGLLVEAGPTTT-SEFSIYIIDEFILIYAPKLIIGSGN 305
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 127 VKCGRGKVDLKKMLDILYDKIKSILLEGCTLWNGMFKGLVDEVSVIAPKIFGCK-- 184
CC
CC Qy 306 YQYOTNDVIEIP---DANQFEIYHSELNQNKLTLR 340
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 185 -----EAPTYVDGEGFKTV-----DECVKLEK 207
CC
CC RESULT 13
CC RIB7_METTH
CC ID RIB7_METTH STANDARD; PRT; 216 AA.
CC AC O26337;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE PUTATIVE 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE
CC DE (EC 1.1.1.193) (HTP REDUCTASE).
CC GN MTH235;
CC OS Methanobacterium thermoautotrophicum.
CC OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
CC OC Methanobacterium.
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-DELTA H;
CC RX MEDLINE; 98037514.
CC RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
CC RA Aldredge T., Bashirzadeh R., Blakely D., Cook K., Gilbert K.,

```

```

RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delat: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000810; AAB84741.1; -.
CC INTERPRO: IPR002734; -.
CC PFAM; PF01872; RibD.C; 1.
CC KW Hypothetical protein; Riboflavin biosynthesis; Oxidoreductase; NADP.
CC SEQUENCE 216 AA; 23576 MW; 436561C089DF41A6 CRC64;
CC
CC Query Match 10.3%; Score 184; DB 1; Length 216;
CC Best Local Similarity 25.6%; Pred. No. 2.5e-07;
CC Matches 53; Conservative 46; Mismatches 98; Indels 10; Gaps 5;
CC
CC Qy 138 POITVKVSASLDGKQANDNGOSQITNKEVKQDVYKLRHRHDAVLGTGRRTVELDDPQVTT 197
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 3 PYVILNAAWTLGDKIATATGSGE-ISGEEDLRRVHLEKRDALWNGTIVLADDPRLTV 61
CC
CC Qy 198 RIQD---GKNPIKVLKSGNIHFNQIYQDESTPIWIYTEN---PNLTS---NQTHTIEI 249
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 62 HRVDAAPGDNPRVVDSMARTPPHFRLVNDAPTIVGSESAAPERVAELKRAEVVA 121
CC
CC Qy 250 YLKSCDLTTLNLYKRGVGLLVEAGPTTT-SEFSIYIIDEFILIYAPKLIIGSGNQVF 308
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 122 GTRRVLDLHLLERLHGMIERLMLEGGSTLNSMLTGGVLVDEVRCIAPMTIVGGRDRTL 181
CC
CC Qy 309 YQYOTNDVIEIPDANQFEIYHSELNQNQV 335
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
CC 182 VDGEIDEMADAIKRLKRSYTLGEDL 208
CC
CC RESULT 14
CC RIB7_SULSO
CC ID RIB7_SULSO STANDARD; PRT; 213 AA.
CC AC P95872;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 01-OCT-2000 (Rel. 40, Last annotation update)
CC DE PUTATIVE 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL REDUCTASE
CC DE (EC 1.1.1.193) (HTP REDUCTASE).
CC GN C06025;
CC OS Sulfolobus solfataricus.
CC OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN-DSM 1617 / P2;
CC RX MEDLINE; 97055432.
CC RA Sensen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,
CC RA Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
CC RA Doolittle W.F., Ragan M.A., Charlebois R.L.;
RT "Organizational characteristics and information content of an
RT archaeal genome: 156 kb of sequence from Sulfolobus solfataricus
RT P2."
RL Mol. Microbiol. 22:175-191(1996).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +

```

```
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- PATHWAY: THIRD STEP OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE HTP REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y08256; CAA69508.1; -.
CC INTERPRO: IPR002734; -.
CC PFAM: PF01872; RibD_C; 1.
CC KW Hypothetical protein; Riboflavin biosynthesis; Oxidoreductase; NADP.
CC SEQUENCE 213 AA; 24561 MW; A741F6CADEF7631C CRC64;
CC -----
Query Match 9.1%; Score 163; DB 1; Length 213;
Best Local Similarity 26.8%; Pred. No. 1e-05;
Matches 57; Conservative 48; Mismatches 92; Indels 16; Gaps 10;
QY 138 PQITVKVSASLDKQANDNCQSWITNKVKQDYKLRHRDAVLTRGVVELDDPQYTT 197
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
DB 5 PYVIFSTVSIDGLATKTKGYSE-LSCPDKQKQHEIRSEVDAMVVGANTVRVDNPSLT 63
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
QY 198 RI-QDGNPKIVILSKSGNIHFNOQIODESTPIWIITNPNTLSNOTHTE--IIYLKSC 254
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
DB 64 KYGKRNRPVRVVRFSNLDPSYKIFTPPSIV-IYTSYESEKVEEFKGVIVRKFL 122
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
QY 255 DLTTILHNLKXR-GVGFTLLVEAGPTTTFSEF---SIYVIDEFLIYAPKLGSGGNYQFYQ 310
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
DB 123 HLDDLEDLDNFRNVRMLWVEGGHLIWFIFKDNLY--DEIRITISPRIFG--NGVSFTQ 178
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
QY 311 TNDVIEIPDANQFEIVHSEL--LNQNVKLFRLK 341
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
DB 179 GEGFIG-EDSPRIELIDAKICEGCEVHLYTKK 210
| : : : | : | : | : : | : | : | : | : | : | : | : | : | : | : |
RESULT 15
YFHC_ECOLI
ID YFHC_ECOLI STANDARD; PRT; 178 AA.
AC F30134;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 20.0 KDA PROTEIN IN PURL-DPJ INTERGENIC REGION (ORF178).
GN YFHC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NWL37;
RX MEDLINE; 92292954.
RA Poulsen L.K., Larsen N.W., Molin S., Andersson P.;
RT "Analysis of an Escherichia coli mutant strain resistant to the cell-
RT killing function encoded by the gef gene family.";
RL Mol. Microbiol. 6:895-905(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Nashimoto H., Saito N.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
```

```
RT The complete genome sequence of Escherichia coli K-12.;
RL Science 277:1453-1474(1997).
CC -1- FUNCTION: MUTATION IN THIS PROTEIN MAKES E.COLI RESISTANT TO THE
CC TOXIC PROTEINS ENCODED BY THE GEF GENE FAMILY.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
CC FAMILY. STRONG, TO E.COLI AND B.SUBTILIS RIBG. STRONG, TO
CC H.INFLUENZAE HI0906.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X72336; CAA51064.1; -.
CC EMBL: D64044; BAA10909.1; -.
CC EMBL: U36841; AAA79821.1; -.
CC EMBL: AE000342; AAC75612.1; -.
CC PIR: S20974; S20974.
CC ECGENE: EG11372; YFHC.
CC INTERPRO: IPR002125; -.
CC PFAM: PF00383; DCMF_Cyt_deam; 1.
CC PROSITE: PS00903; CYT_DCMF_DEAMINASES; 1.
CC KW Hypothetical protein; Hydrolase; zinc.
FT METAL 68 68 ZINC (BY SIMILARITY).
FT METAL 98 98 ZINC (BY SIMILARITY).
FT METAL 101 101 ZINC (BY SIMILARITY).
FT MUTAGEN 64 64 D->E: RESISTANCE TO THE CELL-KILLING
FT FUNCTION ENCODED BY THE GEF GENE FAMILY.
SQ SEQUENCE 178 AA; 20026 MW; 80B2E3B5FD61AA8A CRC64;
-----
Query Match 8.2%; Score 146.5; DB 1; Length 178;
Best Local Similarity 29.3%; Pred. No. 0.00014;
Matches 49; Conservative 29; Mismatches 60; Indels 29; Gaps 10;
QY 1 MDYAIOLANNVGQGTGVNPPGVAVVNEGRIVG-----IGAHLRKGDKHAQVAL---D 51
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 23 MRHALTLAKRWDREY--PVGAVLVHNNRVIGEGWNRPIGRH--DPTAHAEIMALRQGG 78
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 52 MAQQNAE--GATIIITLEPCSHFGSTPPCVNKIIDCKIAKVYVATKDNSLDTHGD--ETL 107
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 79 LVNQVRLIDATLYVTLEPC-----VMCAGAMHSRIGRVVFGARDAKTGAAGSLMDVL 132
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 108 R----AHGTEV-PCVDDERASQLYQDFFRAKAKQLPQITVKVSASLD 149
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 133 HHPGMNHRVEITEGILADECAALLSDFFMRROEI-KAOKKAQSSTD 178
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Search completed: November 14, 2000, 10:10:27
Job time: 176 sec
```

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	530	29.6	348	2	Q9X2B8	Q9x2e8 thermotoga	
2	422	23.6	357	2	Q9RY01	Q9ry01 deinococcus	
3	416	23.3	241	2	Q86961	Q86961 thermotoga	
4	408.5	22.8	375	2	Q84735	Q84735 chlamydia t	
5	396	22.1	245	2	Q51825	Q51825 shigella fl	
6	390.5	21.8	363	2	Q9REF6	Q9ref6 bartonella	
7	320	17.9	599	10	Q9SYV4	Q9sty4 arabidopsis	
8	300	16.8	363	10	Q9SUB7	Q9sub7 arabidopsis	
9	180	10.1	224	2	Q86747	Q86747 aquifex aeo	
10	160	8.9	227	2	Q9REG2	Q9reg2 bartonella	
11	142	7.9	145	2	Q9X3V3	Q9x3v3 pseudomonas	
12	140	7.8	344	2	Q26035	Q26035 helicobacte	
13	140	7.8	344	2	Q9ZJB5	Q9zjb5 helicobacte	
14	139	7.8	164	2	P73717	P73717 synecocyst	
15	134.5	7.5	145	2	Q9RR66	Q9rr66 deinococcus	
16	127.5	7.1	376	2	Q9RKM1	Q9rkm1 streptomyce	
17	125	7.0	1307	10	Q9SY10	Q9sy10 arabidopsis	
18	116.5	6.5	405	3	P87241	P87241 schizosacch	
19	116	6.5	201	2	Q9WYN5	Q9wyn5 thermotoga	

0
9
8
7
6
5
4
3
2
1
0

6 MKRAIELAKKGLGRVNPNPVGAVVVKDGRITIAEGFFHPYFGGPHAEERMAIESAR

Qy 1 MDYAIQLANMVGGQTGYNPPVGVAVWVNEGRIVGIGAHLRKGDKHAEVQALDMAQQNAE-- 58
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Dd 6 MKRAIELAKKGLGRVNPDPVGVAVWVKDGIIEAGPHPYFGGPHAERVAIESARKKGEDL 65

```
QY 59 -GATIIITLEPCSHFGSTPPCVNKKIIDCKIAKVYATKD-NSLDTHGDETLRAHGIEV-E 115
Db 66 RGATLIIVTLEPCDHGKTPTCTDLIIIESGIKTIVTIGTRDPNPVSGNGVEKFRNHGIEVIE 125
QY 116 CVDDERASQLYQDFFRAKAKQLPQITVKVYSASLDGKQANDNGOSOWITNKEVKQDVYKLR 175
Db 126 GVLEEEVKKICEFFITVTKRPFVALKVASTLDGKIADHRGDSKWITDK-LRFKVVHEMR 184
QY 176 HRHDAVLTRGRTEVLDDPOVYTRIOGKNPIKVLKSGNIHFNOQIYO--DESTPIWIY 233
Db 185 NIYSAVLGAGTVLKNPQLTCLKGRNPRVILDRKGVL--SGKVRVFENARVIFV 242
QY 234 TENPNITSNQTHIEIILYKSCDLTTILHNLKRGVCTLLVEAGPTTTSFSIYIIDEFIL 293
Db 243 TESEE-AEYPPHVEKA-LSDCSVESILNLYERDIDSVLVEGGSKVFSEF-LDHADWVFG 299
QY 294 YAPKLLIG-GSGNYQYQYNDVIEIPDANQFEIVHSELLNQNKLTLR 340
Db 300 FYSTKIFGKGLDVFSGYLSD--VSVPP--KFKVNVNFEFSDSEFLVEMR 343

RESULT 2
Q9RY01 PRELIMINARY; PRT; 357 AA.
AC Q9RY01;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE.
GN DR0153.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1.
RX MEDLINE; 20036896.
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL; AF001878; AAF09742.1; -.
DR TIGR; DR0153; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
SQ SEQUENCE 357 AA; 37226 MW; ADBFE65C10626007 CRC64;
```

```
Query Match 23.6%; Score 422; DB 2; Length 357;
Best Local Similarity 36.4%; Pred. No. 1e-25;
Matches 116; Conservative 45; Mismatches 128; Indels 30; Gaps 8;

QY 1 MDYAIQANWVGQOTGVNPPVGAIVVYVNEG-----RIVGIGAHLRKKDKHAEVQALDMAQON 56
Db 12 MQALNEAAKGLRGTSPNPNVPGCVIIVRDGEIAIEIYGRGFHPKAGEPHEAVTALREAGER 71
QY 57 AEGATIIITLEPCSHFGSTPPCVNKKIIDCKIAKVYATKDNS--LDTHGDETLRAHGIEV 114
Db 72 ARGATAYITLEPCSHGORTPPCADALIAAGVAVVVAAGDPNPQVNGRGLKLRAGAGIEV 131
QY 115 ECVDDERASQLYQDFFKA-KAKOLPQITVKVYSASLDGKQANDNGOSOWITNKEVKQDVYK 173
Db 132 ATGVLEAAVROQAGFRSLVTRGRPHVIVKYAMTLDGKVAALNEGNGPVSGPEARARVMA 191
QY 174 LRHRHDAVLTRGRTEVLDDPOYTR-IQDGKNPIKVLKSGNIHFNOQ-----IY 223
```

```
Db 192 WRNEVDVAVGARTALDNPQLNVRLDGGDRPRAVLDFEGHLPASARAVREGTVLVLR 251
QY 224 ODESTPIWIYTENPNITSNQTHIEIILYKSCDLTTILHNLKRGVCTLLVEAGPTTISE- 282
Db 252 EGRSTPL---ERDPRVTVLHAH-----SLQGALEQLAGLGATVLLLEGGPTLASAF 299
QY 283 FSIIYIDEFILYAPKLLIG 301
Db 300 FEAGLIDELRVFVAPKLLG 318

RESULT 3
O86961 PRELIMINARY; PRT; 241 AA.
AC O86961;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE (FRAGMENT).
GN RIBD.
OS Thermotoga neapolitana.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2706-MC24;
RA Zverlov V.V., Mashchenko O.V., Liebl W., Velikodvorskaya G.A.;
RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ009832; CAA08869.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
FT NON_TER 241 241
SQ SEQUENCE 241 AA; 26675 MW; 2C91552C0D897259 CRC64;

Query Match 23.3%; Score 416; DB 2; Length 241;
Best Local Similarity 40.9%; Pred. No. 1.7e-25;
Matches 97; Conservative 46; Mismatches 86; Indels 8; Gaps 6;

QY 1 MDYAIQANWVGQOTGVNPPVGAIVVYVNEGRIYVIGIAH-LRKGDKHAEVQALDMAQONAE- 58
Db 6 MKRAIELAKKGLGRVNPVPPVGAIVVYVKEGRIIEGPHPLFLVGPHEAIVAESAKRKGED 65
QY 59 --GATIIITLEPCSHFGSTPPCVNKKIIDCKIAKVYATKD-NSLDTHGDETLRAHGIE-V 114
Db 66 LTGTTLVVTLPECDHHRGKTPTCTDLIIIESGIKVVIGMRDPNPVSGVGEKLLKHGIEV 125
QY 115 ECVDDERASQLYQDFFKAKAKOLPQITVKVYSASLDGKQANDNGOSOWITNKEVKQDVYK 174
Db 126 EGVLEEVKLLCEFFIYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 184
QY 175 RHRHDAVLTRGRTEVLDDPOYTRIOGKNPIKVLKSGNIHFNOQIYODESTPI 230
Db 185 RNVSVAVLGARTVLKDDPRLTCLKGRNPRVILDRGILSGGSYRVFEDNARVI 241

RESULT 4
O84735 PRELIMINARY; PRT; 375 AA.
AC O84735;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE RIBOFLAVIN DEAMINASE.
GN RIBD.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
```

```

RX MEDLINE; 99000809.
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RL Chlamydia trachomatis";
DL Science 282:754-759(1998);
RE EMBL; AE001343; AAC68325.1; -.
DR INTERPRO; IPR002125; -.
DR DR INTERPRO; IPR002734; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PFAM; PF01872; RibD_C_1
DR PROSITE; PS00903; CYT_DCMF_DEAMINASES; 1.
SQ SEQUENCE 375 AA; 41088 MW; 8E7BF32BD73611E2 CRC64;

Query Match                22.8%; Score 408.5; DB 2; Length 375;
Best Local Similarity      31.8%; Pred. No. 1.3e-24;
Matches 108; Conservative   58; Mismatches 145; Indels    29; Gaps

QY 18 NPPGVAVVNEGRIVGICAHLRKGDHKAHQALDMAQNABGATITYLEPCSHFGSTPP 77
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 29 NPWVCVIVNGCVIGSGHGQGIGSPHAEVCAVDOKCSLEGAEVFVTLECCFHGTRTP 88
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 78 CVNKIIDCKIAKVYA---TKDNSLDTHGETFLRAGHIEVEC-VDERASQLYQDFFKAKA 134
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 89 CDVLILIKSKVAAYVVGLDDPDPVRCKKGVARLAAGIPVYVGVSQEAKTSLSQPPLYQR 148
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 135 KQLPIQTIVKYVASIDLGKDANDGSOMITNKEVKDYKKLRHRDAVLTGRRTVELDDPQ 194
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 149 RGPLPVVMWKTAASLDGGTDARGSSQSWSISGARADYGKLRAESQAII VGARTVCLDNPR 208
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 195 YTFRIQDG----KNPKIIVLKSGNIHFNQIODESTPIWIYNENPLMTSNOTHIELIY 250
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 209 LSARFPHGDIYEROPLRVVDSDRCTVPLESKFV-DLSSGSTLF-----TTQCPEKIY 262
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 251 -LK-----SCDLLTLHLNYKRGVGTLLLVLAGPTTTSEF-SYYVIDFIYYA 296
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 263 KLKDLGVVEWESSHSQVDLKCGLLYRLAEGLCQLVLEGGGAOLHSAFWQQKLNVAGVIWG 322
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 297 PKLI GGSGNFQFYOTNDVIPEDANQFEIVHSELLGNVK 336
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||
QY 323 PKFLGDGOQPMULDLO--LSLVTAEHVRIETSLVRDSVK 360
Db     ||| :||| |::| ::||| :| :||| :||||| ||| |||

RESULT 5
OS1825 PRELMINARY; PRT; 245 AA.
ID OS1825
AC AC OS1825;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
TE TE HOMOLOGUE OF BACILLUS SUBTILIS RIBG.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC OC Shigella.
RN [1]
RP RP SEQUENCE FROM N.A.
RC RC STRAIN=245T;
RL RL Way S.S., Sailustio S., Magliozzo R., Goldberg M.B.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
RD RE EMBL; AF002857; AAB95439.1; -.
DR DR INTERPRO; IPR002125; -.
DR DR INTERPRO; IPR002734; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PFAM; PF01872; RibD_C_1.
DR PROSITE; PS00903; CYT_DCMF_DEAMINASES; 1.
SQ SEQUENCE 245 AA; 27010 MW; 217E26FA4802E072C CRC64;

Query Match                22.1%; Score 396; DB 2; Length 245;
Best Local Similarity      37.6%; Pred. No. 6.7e-24;
Matches 89; Conservative   38; Mismatches 84; Indels    26; Gaps
```

QY 1 MDXAIQLANNVQGOTGYNPPVGAVVYNNEGRIVGIGIAHLRKGDHAEVQALDMAQQNAEGA 60
 I : : | | : : | | | | : | | | : | | | : | | | | : | | | : | |
 Db 7 MARALKLAQRGRFTTHPNPNCVCIVKDGELVGEVGHORAGEPHAEVHALRMAGEKAKGA 66
 QY 61 TIYTILEPCSHFGSTPPCVNKIIDCKIAKVYYATKNSLDTHGDETLR-----AHGI 112
 I : | | | | | | | | | | | | : : | | : | | : | | : | | : | | : | | : | |
 Db 67 TAYVTLEPCSHHGRTGCCDALLAAGVARVVAAMQDPNPQOVARGRLYRLOAQAGIDVSHGL 126
 I : | | | | | | | | | | | | : : | | : | | : | | : | | : | | : | | : | |
 QY 113 EVECDDERASQLYQDFEKKAKOLPOITVKVSASLDGKDANDNGOSOWITNKKEVDYV 172
 : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
 Db 127 MM-----SEPEQLNKGFLKMRTGFPFIQLKLGSLDGRTMAMSGESWITSQAARDVQ 181
 : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
 QY 173 KLRHRHDVLTGRTRVELDDPOQYTTT-----IQDGKNPIKVIKLSKSGNI 216
 : | : | | | | | | | | | | : : | | : | | : | | : | | : | | : | | : | |
 Db 182 RORAQSHAILTSSATVLADDPALTVRWSELDELTQALYPQONLRQPRIRVIDSQNRV 238
 : | : | | | | | | | | | | : : | | : | | : | | : | | : | | : | | : | |

RESULT 6
 Q9REF6 PRELIMINARY; PRT; 363 AA.

ID AC Q9REF6
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE RIBOFLAVIN-SPECIFIC DEAMINASE / REDUCTASE (FRAGMENT).
 GN RIBD.
 OS Bartonella henselae (Rochalimaceae henselae).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bartonellaceae; Bartonella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HOUSTON-1;
 RA Bereswill S., Hinkelmann S., Kist M., Sander A.:
 RT "Molecular analysis of riboflavin synthesis genes in bartonella
 RT henselae and use of the ribC gene for differentiation of bartonella
 RT species by PCR."; J. Clin. Microbiol. 37:3159-3166(1999).
 RL EMBL; AJ132928; CAB63090.1; -;
 DR INTERPRO; IPR000886; -;
 DR INTERPRO; IPR002125; -;
 DR INTERPRO; IPR002734; -;
 DR PFAM; PF00383; dCMP_cyt_deam; 1.
 DR PFAM; PF01872; RibD_C; 1.
 DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
 DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 363 AA; 39524 MW; 131630AF56074F7CD C64;

Query Match 21.8%; Score 390.5; DB 2; Length 363;
 Best Local Similarity 32.7%; Pred. No. 3.2e-23;
 Matches 111; Conservative 60; Mismatches 135; Indels 33; Gaps

QY 4 AIQLANNVQGOTGYNPPVGAVVYNNEGR-----IVGTGAHLRKGDHAEVQALDMAQQNAE 58
 I : | | | | | | | | | | | | : | | | | | | | | | | | | | | | | | | | | |
 Db 2 AIRLAERHVGLTGENTPSVGTIIARNDENVGVYLVGVTAIQGRPHAEVQALMAQS LAH 61
 I : | | | | | | | | | | | | : | | | | | | | | | | | | | | | | | | | | |

QY 59 GATIYITLPCSHFGSTPPCVNKIIDCKIAKVYYATK--DNSLDTHGDETFLRAHIEV-E 115
 I : | | | | | | | | | | | | : | | | | | | | | | | | | | | | | | | | | |
 Db 62 GATAVTVLPCHSYHGKTSPCTNLLKSISRVTALTDLDKRVRNGRIALLRAAGIEVTE 121
 I : | | | | | | | | | | | | : | | | | | | | | | | | | | | | | | | | | |

QY 116 CVDDERASQLYQDFEKKAKOLPOITVKYSASLDGKDANDNGOSOWITNKKEVDYVYLK 172
 I : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
 Db 122 GVLAKEAFESLCNWICIRKOORCAVTLKMAISADNGVGKKGGGINKISGTTSHAQTHILR 181
 I : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |

QY 176 HRHDVLTGRTRVELDDPOQYTTTRI--QDGKNPIKVIKLSKSGNIHFNQOIYQD-ESTPIWI 232
 : : | | | | | | | | | | | | : : | | : | | : | | : | | : | | : | | : | |
 Db 182 AQNNAILVGICTIADDPQLNCRIPGMEMHSPIRIILDANLCIPLANKVVQTAKIPTVW 241
 I : | | | | | | | | | | | | : | | | | | | | | | | | | | | | | | | | | |

QY 233 YTENPNITSNOTHEI-----IYLKSCDLT-----TILHNLYKRGVGTLLVEAGPTTTS 281
 : : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
 Db 242 ICD-VNFSEKSKKIALEOYGVSYSVEVSNGYMPPLTILLOLYRGINSVILLEGVKTGE 300
 I : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |

QY 1 MDXAIQLANNVQGOTGYNPPVGAVVYNNEGRIVGIGIAHLRKGDHAEVQALDMAQQNAEGA 60
 I : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | | : : | |
 Db 7 MARALKLAQRGRFTTHPNPNVCVVKVDGEIVGEGYHQRAEPHAIEVHALRMAGEKAKGA 66
 QY 61 TIYTILEPCSHFGSTPCCVNKIIDCKIAKVYYATKNSLDTHGDETLR-----AHGI 112
 I : | : ||||| | | | | :
 Db 67 TAYVTLEPCSHHGRTGCCDALLAAGVARVVAAMQDPNPQVARGRLYLRQAQAGIDVSHGL 126
 I : | : ||||| | | | | :
 QY 113 EVECDDERASQLYQDFEKKAKOLPOITVKVSASLDGKDANDNGOSOWITNKKEVKQDVY 172
 : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
 Db 127 MM-----SEPEQLNKGFELKMRTGFPFIQLKLGASLDGRTAMASGESWITSQARRDVQ 181
 : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
 QY 173 KLRHRHDVLTGRTRVELDDPOQYTTT-----IQDGKNPIKVIKLSKSGNI 216
 : : | : | | | | | | | | :
 Db 182 RORAQSHAILTSSATVLADDPALTVRWSELDEQTQALYPQONLRQPRIRVIDSQNRV 238
 : : | : | | | | | | | | :
 RESULT 6
 Q9REF6 PRELIMINARY; PRT; 363 AA.
 ID AC Q9REF6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE RIBOFLAVIN-SPECIFIC DEAMINASE / REDUCTASE (FRAGMENT).
 GN RIBD.
 OS Bartonella henselae (Rochalimaceae henselae).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Bartonellaceae; Bartonella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HOUSTON-1;
 RA Bereswill S., Hinkelmann S., Kist M., Sander A.;
 RT "Molecular analysis of riboflavin synthesis genes in bartonella
 RT henselae and use of the rbcC gene for differentiation of bartonella
 RT species by PCR.";
 RL J. Clin. Microbiol. 37:3159-3166(1999).
 DR EMBL: AJ132928; CAB63090.1; -;
 DR INTERPRO: IPR000886; -;
 DR INTERPRO: IPR002125; -;
 DR INTERPRO: IPR002734; -;
 DR PFAM: PF00383; dCMP_cyt_deam; 1.
 DR PFAM: PF01872; RibD_C; 1.
 DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
 DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 363 AA; 39524 MW; 131630AF56074F7CD C64;

Query Match 21.8%; Score 390.5; DB 2; Length 363;
 Best Local Similarity 32.7%; Pred. No. 3.2e-23;
 Matches 111; Conservative 60; Mismatches 135; Indels 33; Gaps

QY 4 AIQLANNVQGOTGYNPPVGAVVYNNEGR-----IVGTGAHLRKGDHAEVQALDMAQQNAE 58
 I : | : | | | | | | | | :
 Db 2 AIRLAERHVGLTGENTPSVGTIIARNDENVGVIVGVTAIQGPRAEVAQALMAQS LAH 61
 I : | : | | | | | | | | :
 QY 59 GATIYITPLEPCSHFGSTPCCVNKIIDCKIAKVYYATK--DNSLDTHGDET LRAH IEV-E 115
 I : | : | | | | | | | | :
 Db 62 GATAVTVTPLEPCSHYGKTSPTCNLLKSGISRVIALTDLDKRVRNGRIALLRAAGIEVTE 121
 I : | : | | | | | | | | :
 QY 116 CVDDERASQLYQDFEKKAKOLPOITVKVSASLDGKDANDNGOSOWITNKKEVKQDYKL 172
 I : | : | : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | :
 Db 122 GVLAKEAFESLCNWICIRKOORCAVTLKMAISADNGVGKKGGGINKISGTTSHAQTHILR 181
 I : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
 QY 176 HRHDVLTGRTRVELDDPOQYTTTRI--QDGKNPIKVIKLSKSGNIHFNQIYQD-ESTPIWI 232
 : : | : | | | | | | | | :
 Db 182 AQNNAILVGICTIADDPQLNCRIPGMEMHSPIRIILDANLCIPLANKVVQTAKIPTVW 241
 I : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
 QY 233 YTENPNITSNOTHEI-----IYLKSCDLT-----TILHNLYKRGVGTLLVEAGPTPTS 281
 : : : : : : | :
 Db 242 ICD-VNFSEKSKKIALEOYGVSVYSVEVSYNGYMPPLTILLOLYRGINSVILLEGVKTGE 300
 : : : : : : | :
 Db 242 ICD-VNFSEKSKKIALEOYGVSVYSVEVSYNGYMPPLTILLOLYRGINSVILLEGVKTGE 300
 : : : : : : | :


```
SQ SEQUENCE 224 AA; 25256 MW; E4A21E2EA4540024 CRC64;

Query Match
Best Local Similarity 10.1%; Score 180; DB 2; Length 224;
Matches 44; Conservative 38; Mismatches 83; Indels 12; Gaps 3;

QY 138 PQITVKVSASLDGQANDNGOS----QWITNKVKQDVYKLRHRHDAVLTRGRIVELDDP 193
| : : : | : | | | | : : : : | : : : | : | | |
Db 4 PYVIIVSEVSDGKLTLRYGASSKELMSLDEAYKYLHEIRAKVDGIMVGCEVTRTNP 63
| : : : | : | | | | : : : : | : : : | : | | |

QY 194 QYTRIOGKNPKIVILSKSNIHFNOQIYODESTPIWIYTN-----PNLTSNQTHI 246
| : : : | : | | | | : : : : | : : : | : | | |
Db 64 SLTVRYAKGNPVRILPCTGANVPLDANVLNKEAPTIIATERAPKRELKIKELGAEV 123

QY 247 EIIYLKSCDLTIIHLNLYKRGVGTLLVEAGPTTTFSEF-SIYVIDEFILYAPKLIKG 302
| : : : | : | | | | : : : : | : : : | : | | |
Db 124 IVVGDELVDKLLPELYRRIKSLMVEGGASINWFEVRRRVVDIEIRLIHLPIVVG 180

RESULT 10
Q9REG2 PRELIMINARY; PRT; 227 AA.
AC Q9REG2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE (FRAGMENT).
GN RIBD.
OS Bartonella clarridgeiae.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bartonellaceae; Bartonella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=73;
RA Bereswill S., Hinkelmann S., Kist M., Sander A.;
RT "Molecular analysis of riboflavin synthesis genes in bartonella
RT henselae and use of the ribC gene for differentiation of bartonella
RT species by PCR.";
RL J. Clin. Microbiol. 37:3159-3166(1999).
DR EMBL: AJ236916; CAB63087.1; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF01872; RibD_C; 1.
FT NON_TER 1
SQ SEQUENCE 227 AA; 25344 MW; CFBB8DFF6A543ED5 CRC64;

Query Match
Best Local Similarity 8.9%; Score 160; DB 2; Length 227;
Matches 59; Conservative 49; Mismatches 87; Indels 38; Gaps 11;

QY 133 KAKQLPQITVKVSASLD---GKQANDNGOSQWITNKVK-ODVYKLRHRHDAVLTRGRTV 188
| : : : | : | | | | : : : : | : : : | : | | |
Db 3 KKLORCEVTLMKATISADNGVCKK-----GQSLKISGEISHSQTHILRAQNDVIWVGICTI 58

QY 189 ELDDPQYTRIQ--DGKNPIKVIILSKSGNIHFNOQIYODEST-PIWIYTNPNLTSNQTH 245
| : : : | : | | | | : : : : | : : : | : | | |
Db 59 LADDPKLDCLRPLGLEMSPRIVLDKLNRIPLDAKVQVQTAANIPTWVIC-GTVLSKRRKK 117

QY 246 IEI----IYLKSCDLT-----TILNLYKRGVGTLLVEAGPTTTFSEF-SIYVIDEIL 293
| : : : | : | | | | : : : : | : : : | : | | |
Db 118 IALPQYGVTVCSVETVNNLIPPLAILRLLYQKRLNSVLEGGAKTGEIFLNASCVDRLIC 177

QY 294 YYAPKLIIGSGNYQFYQTNIDVIEIPD---ANQFEIVHSELLNQNVLTKRKK 342
| : : : | : | | | | : : : : | : : : | : | | |
Db 178 FYAPLIIG-----KDRIEAPHFESYLSFNKTEMRLMGNDRLYKWRKK 220

RESULT 11
Q9X3V3 PRELIMINARY; PRT; 145 AA.
ID Q9X3V3
AC Q9X3V3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
```

```
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CUMB.
GN Pseudomonas putida GB-1.
OS Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
OC Pseudomonas.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-1;
RX MEDLINE: 99203138.
RA Brouwers G.J., Vrind J.P.M., Corstjens P.L.A.M., Cornelis P.,
RA Baysse C., Vrind-de Jong E.W.;
RT "cuma, a gene encoding a multicopper oxidase, is involved in Mn2+
RT oxidation in pseudomonas putida GB-1.";
RL Appl. Environ. Microbiol. 65:1762-1768(1999).
DR EMBL: AF086638; AAD24212.1; -.
DR INTERPRO: IPR002125; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
DR SEQUENCE 145 AA; 15171 MW; A92F44D242DFA70C CRC64;

Query Match
Best Local Similarity 7.9%; Score 142; DB 2; Length 145;
Matches 45; Conservative 24; Mismatches 54; Indels 22; Gaps 6;

QY 5 IQLANWVGQGTGVPVPGAVVNEGRIVGIGAHLRKGDGK----HAEOALDMAQONAE-- 58
| : : : | : | | | | : : : : | : : : | : | | |
Db 3 LALALAAEAGALGEVPGAVLVQGVQGVGQGNRPIDSDPSAHAEMVAIRAAKAAASY 62

QY 59 ---GATYITILEPCSHFGSTPPCVNKKIIDCKIAKVYVAT---KNSLDTHGD---ETLRA 109
| : : : | : | | | | : : : : | : : : | : | | |
Db 63 RLPGSTLYVTLEPCS-----MCAGLIVHSRVMRVFEGALEPKAGIVSQGQFGQGLN 116

QY 110 HGIEVE-CVDDERASQLYQDFFKAK 133
| : : : | : | | | | : : : : | : : : | : | | |
Db 117 HRVIVEGGVLAECGQILSDFFKAR 141

RESULT 12
O26035 PRELIMINARY; PRT; 344 AA.
ID O26035
AC O26035;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN (RIBG).
GN HPI505.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE: 97394467.
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL: AE000648; AAD08541.1; -.
DR TIGR: HPI505; -.
DR INTERPRO: IPR002125; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
KW Hypothetical protein.
```

```

SQ SEQUENCE 344 AA; 39025 MW; 90AEC4D1048191D CRC64;

Query Match
Best Local Similarity 7.8%; Score 140; DB 2; Length 344;
Matches 63; Conservative 37; Mismatches 104; Indels 66; Gaps 10;

QY 18 NPPVGVVNEG-RIVGIGAHLRKDKHAEVQALDMA----- 53
Db 25 NPSVACVLDKNHEILLSLETHKRAKTPHAEVLAQAQSAKILRPSLKNDEKLEDPKTLSD 84
QY 54 -----QONAGATYI-ITLPCSHFGSTPCVNKIIDCKIAKVYATKNSLTHG----- 103
Db 85 FLKTHDHAFTDCVFLITLPCSNYSYKTPACSELLEILKPKRVIIATEENEAKKGGGLARL 144
QY 104 ----DETLRAHGTEVECVDDERASQLYQDFKAKAKOLPOITVKVSASLDGKQANDNGQS 159
Db 145 OKARIETIICHNLE-----NKAKDLLLPRVMEQGRFNL-FKLALRMNGDYHHGK--- 194
QY 160 QWITNKEVKDQVYKLRHRHDAVLTRGRTRVELDDPQYTTRIQD-----GKNPIKVLKSGN 215
Db 195 --ITGQKSVIFTHNQRAICDTLIVSGKTIRTDNPLLDARFCDSFYQKNKPNIALSKR-S 251
QY 216 IHFNQIYQDESTPIWIYTNPNLTSNQTH 245
Db 252 IDPNSKVF-----SAPNRLVNTFH 270

RESULT 13
Q9ZJB5 PRELIMINARY; PRT; 344 AA.
AC Q9ZJB5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PUTATIVE RIBOFLAVIN-SPECIFIC DEAMINASE.
GN RIBD.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 99120557.
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001561; AAD06964.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
SQ SEQUENCE 344 AA; 39072 MW; 189BD1095298477D CRC64;

Query Match
Best Local Similarity 7.8%; Score 140; DB 2; Length 344;
Matches 59; Conservative 32; Mismatches 90; Indels 58; Gaps 8;

QY 18 NPPVGVVNEG-RIVGIGAHLRKDKHAEVQA----- 49
Db 25 NPSVACVLDKNHEILLSLETHKRAKTPHAEVLAQAQSAKILRPSLKNDEKLEDPKTLSD 84
QY 50 -LMAQONA-EGATYITLPCSHFGSTPCVNKIIDCKIAKVYATKNSLTHGDETL 107
Db 85 FLKTRHDNAFKDCVFLITLPCSNYSYKTPACSGILLEILKPKRVIIATEENEAKKGGGLERL 144
QY 108 -RAHGTEVECVDDERASQLYQDFKAKAKOLPOITV-----KVSASLDGKQANDNG 157
Db 145 OKAHIEITIICHNLE-----KAKNLLLPFRVMEQGRFNLFKLALRMNGDYHHGK- 194
```

```

QY 158 QSQWITNKEVKDQVYKLRHRHDAVLTRGRTRVELDDPQYTTRIQD-----GKNPIKVLK 212
Db 195 ----ITGQKSVIFTHNQRAVCDTLIISGKTIRTDNPLLDARFCDSFYHKNKPNIALISK 249

RESULT 14
P73717 PRELIMINARY; PRT; 164 AA.
ID P73717;
AC P73717;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HYPOTHETICAL 18.0 KDA PROTEIN.
GN SLL1631.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97061201.
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res 3:109-136(1996).
DR EMBL; D90909; BAA17764.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASE; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17985 MW; EE03A8329069459D CRC64;

Query Match
Best Local Similarity 7.8%; Score 139; DB 2; Length 164;
Matches 50; Conservative 22; Mismatches 56; Indels 28; Gaps 9;

QY 1 MDYAIQLANNVQGTGVNPPVGVAVVNE-GRIVGIGAHLRKDK-----HAEVALDMA-- 53
Db 12 MQMAIALAEEA-GNVG-EIPVGVAVVNVAMGEVLATGONRKORDQNPATIAHMLAIOTACR 69
QY 54 ---QNAEGATYITLPCSHFGSTPCVNKIIDCKIAKVYATKD---NSLTHGDETL 107
Db 70 RLGHWRNLNCTLTVLEPC-----PWCCTGAILQARLGLLVYGTDPDKGTIDSVFDLAA 123

QY 108 RA-----HGIEVECVDDERASQLYQDFE---KAKAKQ 136
Db 124 SAASNNHILQSLGCVGQEQCREQLQDNWFAQHRARKKQ 159

RESULT 15
Q9RR66 PRELIMINARY; PRT; 145 AA.
ID Q9RR66;
AC Q9RR66;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE DEOXYCYTIDYLATE DEAMINASE, PUTATIVE.
GN DR2631.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=R1;
RX MEDLINE; 20036896.
RA White O., Eise J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Dally M.J., Minton K.W., Fleischmann R.D.,
```


RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.W.;
RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE002092; AAF12167.1; -.
DR TIGR: DR2631; -.
DR INTERPRO: IPR002125; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
SQ SEQUENCE 145 AA; 15547 MW; 110CA05197EE999A CRC64;

Query Match 7.5%; Score 134.5; DB 2; Length 145;
Best Local Similarity 35.3%; Pred. No. 0.0016;
Matches 42; Conservative 17; Mismatches 29; Indels 31; Gaps 6;
QY 21 VGAVVVN-EGRIVGIGALR-----KGDKHAEOALDMAQOQNAEGATYI 64
Db |||:::|:|:|:|:| ||||| || | ||| |:|:
28 VGACILDRHVRVGVGVNGRAAGPNERSLAOGASGYIHAENVALLAANWNGEGHTLY 87
QY 65 TLEPCSHFGSTPPCVNKKI DC-KIAKVYVATK-----DNSLDTHGDETLRAHGIEVE 115
Db ||||| | :|:|:|:| | :|:|:|:| | :|:|:|:|
88 THEPCS-----VCARLIVNSRRVGVVFATPYRETARVEAGLPS-GAEILRSAGIEVE 139

Search completed: November 14, 2000, 10:10:08
Job time: 205 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:07:20 ; Search time 28.37 Seconds
(without alignments)
119.323 Million cell updates/sec

Title: US-08-978-456-4

Perfect score: 498

Sequence: 1 MDYATQLPNWVGXTGVNPP.....NKIIDCKTAXVVLXNXXRQFR 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	98-0	99	18 W28084	Amino acid sequenc
2	488	98-0	99	20 Y15914	Pyrimidine deamina
3	451	90-6	342	20 Y15913	Pyrimidine deamina
4	261	52.4	361	12 R10070	B.subtilis rib ORF
5	261	52.4	1855	20 Y21803	B. subtilis rib op
6	261	52.4	1855	21 Y83271	Polypeptide encode
7	215	43.2	366	20 Y17949	S. pneumoniae ribc
8	203	40.8	401	21 Y96284	Sinorhizobium meli
9	195	39.2	170	20 Y37428	Amino acid sequenc
10	193	38.8	376	20 Y35519	Chlamydia pneumoni
11	180.5	36.2	410	19 W55031	Actinobacillus ple
12	180.5	36.2	410	20 Y21972	APP ribg gene prod

13	102.5	20.6	155	19	Y86119	S. pneumoniae deri
14	97	19.5	609	16	R88279	Ashbya gossypii DR
15	83.5	16.8	178	18	W18205	Human deoxycytidyl
16	83	16.7	344	19	W98756	H. pylori GHPO 107
17	78	15.7	591	15	R52825	DRAP-deaminase. S
18	76	15.3	158	11	R08285	Thermally stable c
19	76	15.3	158	16	R66507	Yeast cytosine-dea
20	76	15.3	158	21	Y55104	Chimeric cytosine
21	76	15.3	373	20	Y53913	Amino acid sequenc
22	75.5	15.2	163	20	Y35657	Chlamydia pneumoni
23	73	14.7	158	21	Y55101	Chimeric cytosine
24	73	14.7	158	21	Y55103	Chimeric cytosine
25	69.5	14.0	176	20	Y37278	Amino acid sequenc
26	68	13.7	158	21	Y55102	Chimeric cytosine
27	66	13.3	569	16	R67378	H. pylori urease u
28	62.5	12.6	155	19	R38699	S. pneumoniae come
29	62	12.4	569	11	R04580	Part of protein w
30	62	12.4	569	17	W07194	H. pylori urease B
31	62	12.4	569	19	W98511	H. pylori GHPO 124
32	61	12.2	559	12	R13550	B subunit of H. py
33	61	12.2	749	19	W80599	Helicobacter pylor
34	61	12.2	2441	16	R79054	CREB binding prote
35	61	12.2	2441	19	W40058	Cellular transcrip
36	61	12.2	2441	21	Y94252	Mouse nuclear CREB
37	60.5	12.1	169	19	W37847	Human endokine-alp
38	60.5	12.1	169	21	Y53061	Human endokine alp
39	60.5	12.1	177	20	Y06646	Human PRO364 ligan
40	60.5	12.1	177	20	Y15817	Amino acid sequenc
41	60	12.0	346	21	Y68647	Amino acid sequenc
42	58.5	11.7	405	20	Y05469	Human CBP protein
43	58	11.6	145	19	W60119	Mycobacterium vacc
44	58	11.6	145	20	Y14865	M. vaccae antigen
45	58	11.6	273	19	W60128	M. vaccae antigen

ALIGNMENTS

RESULT: 1
ID W28084 standard; Protein; 99 AA.
XX W28084;
AC W28084;
XX

01-SEP-1998 (first entry)
XX

Amino acid sequence of riboflavin biosynthesis protein RIBG.
XX

Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome.
XX

Staphylococcus aureus.
OS

Key Location/Qualifiers
XX

Misc-difference 1..99 /note= "residues designated X are not defined in the specification"
FT
FT

W09730070-AL.
XX

21-AUG-1997.
XX

19-FEB-1997; 97WO-US02318.
XX

20-FEB-1996; 96US-0011888.
XX

(SMIK) SMITHKLINE BEECHAM CORP.
XX

Black MT, Burham MK, Hodgson JE, Knowles DJC, Nicholas RO;
PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
XX


```

XX      SQ      Sequence      361 AA;

Query Match      52.4%;      Score 261;      DB 12;      Length 361;
Best Local Similarity 56.7%;      Pred. No. 2.9e-25;
Matches 51;      Conservative 12;      Mismatches 27;      Indels 0;      Gaps 0;

QY      1      MDYAIIQLPNWVGXTGYNPPYCAVNVNCRIVGICAHLRKGDKHAEVQALDMAQXNAEGA 60
      |||      :||      |||      |||      |||      |||      |||      |||      |||      |||
Db      6      mklaIdAtakgeqctsnPlvgvvvdqglvgmgahLkygeahevhaihmagahea 65

QY      61      TIYTILEPCSHFGSTPPCVNKKIIDCKIAXV 90
      |||      |||      |||      |||      |||      |||      |||      |||      |||
Db      66      diYvtlEpcshYgktppcaeiIlnsgikrv 95

RESULT      5
Y21803
ID      Y21803 standard: Protein. 1855 AA

```

AC	YZ1803;
XX	
DT	10-SEP-1999 (first entry)
XX	
DE	B. subtilis rib operon protein translated from reading frame 3.
XX	
KW	Riboflavin; open reading frame; ORF; structural gene; promoter;
KW	vitamin B2; Bacillus subtilis; rib operon.
XX	
OS	Bacillus subtilis.
XX	
PH	Key Location/Qualifiers

XX /note= "residues xaa are encoded by stop codons"
 PN US5925538-A.
 PD 20-JUL-1999.
 XX
 XX 24-AUG-1998; 980S-0138775.
 XX
 XX 11-SEP-1990; 90US-0581048.
 PR 22-JUN-1989; 89US-0370378.
 PR 21-APR-1992; 92US-0873572.
 PR 06-FEB-1995; 95US-0384626.
 PR 24-AUG-1998; 98US-0138775.
 XX
 XX (HOFF) ROCHE VITAMINS INC.
 XX
 PI Erdenberger T, Hatch RT, Hermann T, Perkins JB;
 PI Pero JG, Sloma A;
 XX
 XX WPI: 1999-418271/35.
 DR N-PSDB; X81947.
 XX
 XX Use of bacterial strains for the over production of riboflavin
 PT
 XX Example 2; Fig 3A-S; 61pp; English.
 PS
 XX The invention relates to a method of producing riboflavin that comprises,
 CC providing a recombinant bacterium containing an exogenously introduced
 CC bacterial ORF (open reading frame) 3 (or ORF5) structural gene and an
 CC exogenous promoter, culturing the bacterium, and recovery of the product.
 CC The method is useful for the production of large quantities of riboflavin
 CC (vitamin B2). Sequences Y21801-803 are protein sequences translated from
 CC three different reading frames of B. subtilis rib operon. Y21804-806 are
 CC protein sequences translated from three different reading frames of the
 CC complementary sequence.
 XX
 XX Sequence 1855 AA;
 SQ

PI Perkins JB, Pero JG, Sloma A;
XX WPI; 2000-320439/28.
DR N-PSDB; Z93752.
XX
XX
PT Novel vector, used for the large scale production of riboflavin,
PT comprises a bacterial or yeast riboflavin biosynthetic protein under
PT the control of a foreign transcription element
XX
XX Disclosure; Figure 3; 66pp; English.
XX
XX Vectors comprising a nucleic acid sequence of bacterial or
XX yeast origin, coding for one or more riboflavin biosynthetic proteins,
XX and one or more transcription elements not naturally associated with
XX the nucleic acid sequence are new. The vector and bacteria transformed
XX with it are used for the large scale production (over 10g/l) of
XX riboflavin. The riboflavin produced can be used to treat disorders
XX associated with riboflavin deficiency such as hair loss, skin
XX inflammation, vision deterioration and growth failure. This
XX polypeptide is a translation of one of the three reading frames of
XX the rib operon sequence described in GENESEQ record Z93752. See
XX also Y83269-74.
XX
SQ Sequence 1855 AA;

Query Match 52.4%; Score 261; DB 21; Length 1855;
Best Local Similarity 56.7%; Pred. No. 2.4e-24;
Matches 51; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

QY 1 MDYAIQLPNMVQGTGVNPPVGVVNEGRIVGTGAHLRKGDKHAEVQALDMAOXNAEGA 60
Db 372 mkaldlakqdegqesnplvgavvkdqgvgmgahlkysgeahvhalhmagahaega 431
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXV 90
Db 432 diyvtlpcshygtktpcaeliinsgikrv 461

RESULT 7
Y17949
ID Y17949 standard; Protein; 366 AA.
AC Y17949;
XX
XX 04-AUG-1999 (first entry)
XX
XX S. pneumoniae ribG polypeptide.
XX
XX Streptococcus pneumoniae; pyrimidine deaminase; pyrimidine reductase;
KW ribG; immune response; antimicrobial; H. pylori infection; cancer;
KW gastrointestinal carcinoma; gastric ulcer; gastritis.
XX
XX Streptococcus pneumoniae.
XX
XX W09927126-A1.
XX
XX 03-JUN-1999.
XX
XX 23-NOV-1998; 98WO-US25010.
XX
XX 25-NOV-1997; 97US-0979616.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX Black MT, Debouck C, Fedon JC, Hodgson JE, Jaworski DD;
PI Knowles DJC, Kosmatka AL, Lonetto MA, Mooney J, Nicholas RO;
PI Palmer LM, Shilling LK, Stodola RK, Wang M, Warren RL;
PI Zhong Y;
XX
XX WPI; 1999-347726/29.
DR N-PSDB; X77179.
XX

PT New Streptococcus pneumoniae ribG polypeptide and polynucleotide
XX useful in the treatment of gastric ulcer and gastritis
XX
XX Claim 14; Page 6; 48pp; English.
XX
XX The invention relates to Streptococcus pneumoniae pyrimidine deaminase
XX and pyrimidine reductase (ribG). S. pneumoniae ribG and its antagonists a
XX re used to treat individuals in need of the polypeptide. Disease related
XX to expression or activity of ribG can be determined by analysing the
XX nucleic acid sequence encoding ribG or detecting the ribG polypeptide in
XX a sample. ribG can also be used to identify antagonists or agonists.
XX ribG, or its related nucleic acid, also has use as a vaccine to induce
XX an immunological response in an animal. Antimicrobial compounds (e.g.
XX agonists and antagonists of ribG), especially broad-spectrum antibiotics,
XX may be of use in the treatment of Helicobacter pylori infection. This
XX should decrease the advent of H. pylori-induced cancers, such as
XX gastrointestinal carcinoma. The treatment should also cure gastric ulcers
XX and gastritis. The present sequence represents the S. pneumoniae ribG
XX polypeptide.
XX
SQ Sequence 366 AA;

Query Match 43.2%; Score 215; DB 20; Length 366;
Best Local Similarity 46.7%; Pred. No. 2.2e-19;
Matches 43; Conservative 12; Mismatches 37; Indels 0; Gaps 0;

QY 1 MDYAIQLPNMVQGTGVNPPVGVVNEGRIVGTGAHLRKGDKHAEVQALDMAOXNAEGA 60
Db 7 mklalkaqgagvynpvmvgalivkdhlihgqgyheffggphaernalknckrspvga 66
QY 61 TIYITLPCSHFGSTPPCVNKKIIDCKIAXVVL 92
Db 67 tlyvtlpcshfghtktpcaidalsgitrvi 98

RESULT 8
Y96284
ID Y96284 standard; Protein; 401 AA.
AC Y96284;
XX
XX 11-AUG-2000 (first entry)
XX
XX Sinorhizobium meliloti RibD.
XX
XX Plant growth; photosynthesis; alfalfa; riboflavin synthase;
KW riboflavin synthesis; RibD.
XX
XX Sinorhizobium meliloti.
XX
XX W0200029607-A1.
XX
XX 25-MAY-2000.
XX
XX 17-NOV-1999; 99WO-US27318.
XX
XX 17-NOV-1998; 98US-0193600.
XX
XX 17-NOV-1998; 98US-0193801.
XX
XX 06-AUG-1999; 99US-0369955.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Phillips DA, Joseph CM, Sanborn JR, Yang G;
XX
XX WPI; 2000-387815/33.
XX
XX N-PSDB; A27366.
XX
XX Enhancing plant growth by applying an agent comprising lumichrome or
XX lumichrome-releasing microorganism or growing plants in hydroponic
XX culture system or in a medium comprising lumichrome-releasing
XX microorganism
XX

PS Claim 61; Page 80-81; 86pp; English.

XX The present sequence is the *Sinorhizobium meliloti* RibD protein,
CC which is involved in riboflavin synthesis. Riboflavin is important
CC in plant growth and photosynthesis, and the protein and gene encoding it
CC can be used to create genetically engineered microorganisms which produce
CC an increased amount of the compound, which in turn increases plant
CC growth. In addition to being applied to the growing plant, the bacteria
CC can also be applied to the seed before planting. The bacteria can be
CC applied to any plant and any part of the plant.

XX Sequence 401 AA;

Query Match 40.8%; Score 203; DB 21; Length 401;

Best Local Similarity 45.7%; Pred. No. 8.4e-18;

Matches 42; Conservative 10; Mismatches 40; Indels 0; Gaps 0;

QY 1 MDYAIQLPNMVQXGTGVNPPVGVAVVYVNEGRIYIGAHLRKGDHAEVQALDMAQXNAEGA 60

Db 12 maaalrlarrnlgltstnpsvgcivnkgitvgravtapggrphaetqalaeagekarga 71

QY 61 TIVITLPCSHFGSTPPCVNKKIIDCKIAWVL 92

Db 72 tayvalepcshhgktppcadaliagvgrvv 103

RESULT 9

Y37428

ID Y37428 standard; Protein; 170 AA.

XX AC Y37428;

XX DT

XX DT

XX 07-OCT-1999 (first entry)

DE Amino acid sequence of a *Chlamydia trachomatis* protein.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.

XX *Chlamydia trachomatis*.

XX WO9928475-A2.

XX PN

XX PD

XX 10-JUN-1999.

XX PF

XX 27-NOV-1998; 98WO-IB01939.

XX PR

XX 04-NOV-1998; 98US-0107077.

XX PR

XX 28-NOV-1997; 97FR-0015041.

XX PR

XX 17-DEC-1997; 97FR-0016034.

XX XX

XX (GEST) GENSET.

XX PA

XX Griffais R;

XX PI

XX WPI; 1999-371125/31.

XX DR

XX Genome sequence of *Chlamydia trachomatis*

XX PT

XX Disclosure; Page 1127-1128; 1755pp; English.

XX PS

XX Y36754-Y37949 are encoded by open reading frames (ORFs) of the genome of

CC *Chlamydia trachomatis* (see 201425). The polypeptides can be used as

CC vaccines against *Chlamydia trachomatis*. Antisense and ribozyme sequences

CC can also be used to control growth of the microorganism. *Chlamydia*

CC *trachomatis* is responsible for a large number of diseases, e.g. eye

CC diseases such as conventional trachoma, nonendemic trachoma,

CC paratrachoma, and inclusion conjunctivitis; genital diseases such as

CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,

CC perihhepatitis, bartholinitis; pneumopathy in breast feeding infants;

CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.

XX Sequence 170 AA;

Query Match 39.2%; Score 195; DB 20; Length 170;

Best Local Similarity 50.7%; Pred. No. 2.9e-17;

Matches 37; Conservative 11; Mismatches 25; Indels 0; Gaps 0;

QY 18 NPPVGAVVYVNEGRIYIGAHLRKGDHAEVQALDMAQXNAEGATIVITLPCSHFGSTPP 77

Db 69 npwvgcivlknvgcivgwhgigsphaevcaqvdkcslegaevyvtilepcchfgrtpp 128

QY 78 CVNKKIIDCKIAWV 90

Db 129 cvdlliikskaav 141

RESULT 10

Y35519

ID Y35519 standard; Protein; 376 AA.

XX AC Y35519;

XX DT

XX 13-SEP-1999 (first entry)

XX DE

XX *Chlamydia pneumoniae* transmembrane protein sequence.

XX KW

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;

XX KW

XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

XX KW

XX vaccine; neutralising epitope.

XX OS

XX *Chlamydia pneumoniae*.

XX PN

XX WO9927105-A2.

XX PD

XX 03-JUN-1999.

XX PF

XX 20-NOV-1998; 98WO-IB01890.

XX XX

XX 04-NOV-1998; 98US-0107078.

XX PR

XX 21-NOV-1997; 97FR-0014673.

XX XX

XX (GEST) GENSET.

XX XX

XX Griffais R;

XX PI

XX WPI; 1999-357842/30.

XX DR

XX Genome sequence of *Chlamydia pneumoniae*

XX PS

XX Page 1273-1274; Disclosure; 1912pp; English.

XX CC

XX Y34584-Y35879 represent the proteins encoded by all the open reading

CC frames in the complete genome (see X91990) of *Chlamydia pneumoniae*.

CC C. pneumoniae causes respiratory disease such as pneumonia and

CC bronchitis and is thought to be a contributing factor in heart

CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema

CC nodosum or pharyngitis. The polypeptides encoded by the open reading

CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in

CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae

CC nucleotide sequences can also be used as immunogenic compositions,

CC especially where the vector directs the expression of a neutralising

CC epitope of C. pneumoniae.

XX Sequence 376 AA;

Query Match 38.8%; Score 193; DB 20; Length 376;

Best Local Similarity 50.7%; Pred. No. 1.5e-16;

Matches 37; Conservative 9; Mismatches 27; Indels 0; Gaps 0;


```

XX DE S. pneumoniae derived protein #328.
XX KW Treatment; prevention; disease; diagnosis; gene therapy; screening;
XX KW bacterial; antimicrobial; antibiotic; pathogenesis; infection.
XX OS Streptococcus pneumoniae.
XX PN WO9806734-A1.
XX PD 19-FEB-1998.
XX PF 15-AUG-1997; 97WO-US14436.
XX PR 16-AUG-1996; 96US-0024022.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
XX PI Stodola RK;
XX WPI: 1998-159452/14.
XX DR N-PSDB; 296435.
XX ST Streptococcus pneumoniae proteins and related DNA - useful for
XX PT screening compounds for antibacterial activity
XX PS Claim 5; Page 589; 640pp; English.
XX CC This invention describes novel isolated Streptococcus pneumoniae
XX CC polynucleotides (see 296173-296494) and their encoded proteins (see
XX CC Y85792-Y86182). The DNA, vectors and host cells described in the method
XX CC of the invention are useful for the recombinant expression of the
XX CC polypeptides. The polypeptides are useful for treatment or prevention of
XX CC disease, or diagnosis of disease related to expression or activity of
XX CC such a polypeptide. They can also be used to screen for compounds which
XX CC interact with and inhibit or activate such a polypeptide. The
XX CC polypeptides (or DNA encoding them, via gene therapy) are also useful
XX CC for inducing an immunological response in a mammal. The antagonists are
XX CC useful to inhibit such bacterial polypeptides. The polypeptides are
XX CC particularly useful to identify antimicrobial compounds and antibiotics.
XX CC They are also useful to determine their role in pathogenesis of
XX CC infection, dysfunction and disease.
XX CC Sequence 155 AA;
XX QY 20 PVGAVVNVNIGRIVGIG---AHLRKGDKHAEOALDMAOXNAEG-----ATTIVITLPC 69
XX DB 29 pigcvivkdgeiigrhnaeelqgravhmaelmaledanlseeswrlldctifvltiepc 87

RESULT 14
R88279
ID R88279 standard; Protein; 609 AA.
XX AC R88279;
XX DT 14-MAY-1996 (first entry)
XX DE Ashbya gossypii DRAP-deaminase encoded by the rib 2 gene.
XX KW Riboflavin biosynthesis pathway enzyme; fungus; Ashbya gossypii;
XX KW complementation; S.cerevisiae; mutant; eukaryote;
XX KW GTP-cyclohydrolase; DRAP-deaminase; DBP-synthase; DMRL-synthase;
XX KW riboflavin synthase; HTP-reductase.
XX OS Ashbya gossypii.

Query Match 20.6%; Score 102.5; DB 19; Length 155;
Best Local Similarity 33.9%; Pred. No. 1.7e-05;
Matches 20; Conservative 14; Mismatches 16; Indels 9; Gaps 2;

QY 20 PVGAVVNVNIGRIVGIG---AHLRKGDKHAEOALDMAOXNAEG-----ATTIVITLPC 69
DB 29 pigcvivkdgeiigrhnaeelqgravhmaelmaledanlseeswrlldctifvltiepc 87

RESULT 15
W18205
ID W18205 standard; Protein; 178 AA.
XX AC W18205;
XX DT 20-AUG-1997 (first entry)
XX DE Human deoxycytidylate deaminase.
XX KW Recombinant deaminase; dCMP.
XX OS Homo sapiens.
XX PN US5622851-A.
XX PD 22-APR-1997.
XX PF 10-JAN-1995; 95US-0370975.
XX PR 10-JAN-1995; 95US-0370975.
XX PA (HEAL-) HEALTH RES INC.
XX PI Maley F, Maley GR, Weiner KXB;
XX DR WPI; 1997-244391/22.

Query Match 19.5%; Score 97; DB 16; Length 609;
Best Local Similarity 38.0%; Pred. No. 0.0005;
Matches 30; Conservative 8; Mismatches 27; Indels 14; Gaps 5;

QY 19 PP-----VGVVNVNIGRIVGIG-AHLRKGDKHAEOALDMAOXNAE----GATYIT 65
DB 477 ppaktsfsgvavlvngteilatgysrelegnthaeqcalqkyfeqhtkdkvpigtviytt 536

QY 66 LEPGS-HFGSTPPCVNKKII 83
DB 537 mepcslrlsgnkpoverii 555

RESULT 15
W18205
ID W18205 standard; Protein; 178 AA.
XX AC W18205;
XX DT 20-AUG-1997 (first entry)
XX DE Human deoxycytidylate deaminase.
XX KW Recombinant deaminase; dCMP.
XX OS Homo sapiens.
XX PN US5622851-A.
XX PD 22-APR-1997.
XX PF 10-JAN-1995; 95US-0370975.
XX PR 10-JAN-1995; 95US-0370975.
XX PA (HEAL-) HEALTH RES INC.
XX PI Maley F, Maley GR, Weiner KXB;
XX DR WPI; 1997-244391/22.

PN DE4420785-A1.
XX PD 05-OCT-1995.
XX PF 15-JUN-1994; 94DE-4420785.
XX PR 25-MAR-1994; 94DE-4410382.
XX PA (BADI ) BASF AG.
XX PI Buitrago SERNA MJ, Revuelta DOVAL JL, Santos GARCIA MA;
XX WPI: 1995-345658/45.
XX DR N-PSDB; T03515.
XX ST DNA encoding riboflavin biosynthesis enzymes - isolated from Ashbya
XX PT gossypii and useful for prodn. of riboflavin in eukaryote hosts
XX PS Claim 2; Page 11-14; 33pp; German.
XX CC The amino acid sequence of the riboflavin biosynthesis pathway enzyme
XX CC DRAP-deaminase encoded by the rib 2 gene from the fungus Ashbya gossypii
XX CC (DRAP is 2,5-diamino-6-ribosylamino-4-(3H)-pyridine-5-phosphate).
XX CC The gene comprises a 450 bp 5'-untranslated region (UTR) contg.
XX CC regulatory sequences, a 1830 bp coding sequence followed by a 347 bp
XX CC 3'-UTR. The gene was isolated from an A.gossypii cDNA library and clones
XX CC were isolated by functional complementation in S.cerevisiae mutants.
XX CC The riboflavin genes (T03514-19) can be used for prodn. of riboflavin in
XX CC eukaryotic cells.
XX CC Sequence 609 AA;

```

DR	N-PSDB; T71696, T71702.
XX	
PT	DNA encoding human deoxycytidylate deaminase - for production of recombinant deaminase
PT	
XX	
PS	Disclosure; Column 109-110; 58pp; English.
XX	
CC	The present sequence represents human deoxycytidylate (dCMP) deaminase. The gene for dCMP deaminase contains a 5' untranslated region (including the promoter), 5 exons, 4 introns and a 3' untranslated region (including the stop signals). The gene can be used to produce recombinant dCMP deaminase, which can be used to convert dCMP to dUMP. Also, the dCMP gene can be altered (removed or mutated) to alter DNA replication in cells, which may lead to mutagenesis.
CC	
XX	
SQ	Sequence 178 AA;
	Query Match 16.8%; Score 83.5; DB 18; Length 178;
	Best Local Similarity 26.9%; Pred. No. 0.0053;
	Matches 28; Conservative 15; Mismatches 30; Indels 31; Gaps 5;
QY	18 NPPVGAYVVN-EGRIVGIG-----AHLRKGDK-----HAEQQA-IDM 52 : : : : :
Db	34 nsqvgacivnsenkigvinyngmgcsddvlpwrrtaenkltdkypvychaeinaimnk 93 : : : : :
QY	53 AQXNAEGATITILEPCSHFGSTPPCVNKIIDCKIAVVLNXR 96 : : : : : : : : : : : : :
Db	94 nstdvkgcsmvalfpncne-----cakiilgagikeviftsdK 131 : : : : : : : : : : : : :

Search completed: November 14, 2000, 10:07:25
Job time: 43 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:07:45 ; Search time 23.2 Seconds
(without alignments)
71.524 Million cell updates/sec

Title: US-08-978-456-4
Perfect score: 498
Sequence: 1 MDYAIQLPNVVGXGTVNPP.....NKIIDCKTAXVVLNXXRQFR 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues
Total number of hits satisfying chosen parameters: 164575

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	98.0	99	3	US-08-978-456-4
2	451	90.6	342	3	US-08-978-456-2
3	180.5	36.2	410	2	US-08-741-327E-15
4	97	19.5	609	2	US-08-716-301-4
5	83.5	16.8	178	1	US-08-370-975B-14
6	76	15.3	155	5	5338678-1
7	76	15.3	158	5	5338678-2
8	66	13.3	569	2	US-08-467-822-26
9	62	12.4	566	2	US-08-920-095-3
10	62	12.4	566	4	PCT-US96-05800-3
11	61	12.2	2441	1	US-08-194-468-2
12	61	12.2	2441	3	US-08-961-739-2
13	60.5	12.1	169	2	US-08-912-227-2
14	58	11.6	145	2	US-08-997-080-55
15	58	11.6	145	2	US-08-997-362-55
16	58	11.6	145	3	US-08-873-970-55
17	58	11.6	273	2	US-08-997-080-75
18	58	11.6	273	2	US-08-997-362-75
19	58	11.6	273	3	US-08-873-970-75
20	58	11.6	370	2	US-08-997-080-194
21	58	11.6	370	2	US-08-997-362-194
22	57.5	11.5	214	1	US-08-277-231A-14
23	57.5	11.5	214	2	US-08-473-750-3
24	57.5	11.5	214	2	US-08-477-326-3
25	56.5	11.3	638	2	US-08-422-699A-11
26	56.5	11.3	638	2	US-08-422-706B-11
27	55.5	11.1	443	2	US-08-620-605D-3
28	55.5	11.1	443	2	US-09-005-232A-3

29	55.5	11.1	448	3	US-08-476-509B-2	Sequence 2, Appli
30	55.5	11.1	486	3	US-08-348-518C-2	Sequence 2, Appli
31	55.5	11.1	555	3	US-08-581-148C-16	Sequence 16, Appl
32	55.5	11.1	1525	3	US-09-191-647-2	Sequence 2, Appli
33	55.5	11.1	1872	5	5386025-6	Patent No. 5386025
34	55.5	11.1	1873	1	US-08-435-675B-4	Sequence 4, Appli
35	55.5	11.1	1873	1	US-08-336-257A-7	Sequence 7, Appli
36	55.5	11.1	1940	2	US-08-644-271-30	Sequence 30, Appl
37	55	11.0	569	2	US-08-467-822-21	Sequence 21, Appl
38	54.5	10.9	338	2	US-08-784-651-6	Sequence 6, Appli
39	53.5	10.7	1996	2	US-08-804-227C-9	Sequence 9, Appli
40	53.5	10.7	1996	2	US-08-804-198-3	Sequence 3, Appli
41	52.5	10.5	164	2	US-08-824-405-2	Sequence 2, Appli
42	52.5	10.5	252	2	US-08-824-405-4	Sequence 4, Appli
43	52.5	10.5	336	2	US-08-784-651-4	Sequence 4, Appli
44	52	10.4	190	1	US-07-956-700B-15	Sequence 15, Appl
45	52	10.4	190	1	US-07-956-700B-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-08-978-456-4
; Sequence 4, Application US/08978456
; Patent No. 6010881
; GENERAL INFORMATION:
; APPLICANT: Palmer, Leslie M.
; TITLE OF INVENTION: No. 6010881el ribg
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,456
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-978-456-4

Query Match 98.0%; Score 488; DB 3; Length 99;
Best Local Similarity 100.0%; Pred. No. 5.2e-59;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDYAIQLPNVVGXGTVNPPGVAVVNEGRIVGIGIAHLRKGDKHAEVQALDMAQXNAEGA 60
Db 1 MDYAIQLPNVVGXGTVNPPGVAVVNEGRIVGIGIAHLRKGDKHAEVQALDMAQXNAEGA 60

```

: ADDRESSEE: Keil & Weinkauff
:
: STREET: 1101 Connecticut Avenue
:
: CITY: Washington
:
: STATE: D.C.
:
: COUNTRY: USA
:
: ZIP: 20036
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
:
: COMPUTER: IBM compatible
:
: OPERATING SYSTEM: MS-DOS version 7.0
:
: SOFTWARE: Wordperfect version 5.1
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/716, 301
:
: FILING DATE: 24-SEP-1996
:
: CLASSIFICATION: 435
:
: CLASSIFICATION: C 12 N 15/53
:
: CLASSIFICATION: C 12 N 15/54
:
: CLASSIFICATION: C 12 N 15/55

```


NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05800
FILING DATE: 23-APR-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,041
FILING DATE: 28-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/568,122
FILING DATE: 06-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06137/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 566 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-05800-3

Query Match 12.4%; Score 62; DB 4; Length 566;
Best Local Similarity 22.1%; Pred. No. 4;
Matches 21; Conservative 14; Mismatches 40; Indels 20; Gaps 3;
Qy 3 YAIOLPNMVGXGTGVNPPVGVVNEGR-----IVGIGAHLRKGDKHAQVQA 49
Db 149 FASGVTTMIGGCTGPDCTNATTITPGRNLKWLRAAEYSMLG-FLAKGNASNDASL 207
Qy 50 LDMAQXNAEGATITILEPCSHFGSTPPCVNKIID 84
Db 208 ADQIEAGAIGFKIH-----EDWGTTPSAINHLD 236

RESULT 11
us-08-194-468-2
Sequence 2, Application US/08194468
Patent No. 5750336
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: ASSAYS FOR THE IDENTIFICATION OF
TITLE OF INVENTION: COMPOUNDS WHICH INHIBIT ACTIVATION OF CAMP AND MITOGEN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/194,468
FILING DATE: 10-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9672
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)-546-4737
TELEFAX: (619)-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2441 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-194-468-2

Query Match 12.2%; Score 61; DB 1; Length 2441;
Best Local Similarity 25.3%; Pred. No. 40;
Matches 21; Conservative 14; Mismatches 32; Indels 16; Gaps 3;
Qy 6 OLPNMVGXGTG--VNPPVGVVNEGRIVGIGAHLRKGDKHAQVQALDMAQXNAEG---- 59
Db 68 QLSLLRGSGSSINPGIGNVSASSPVQOGLG-----GQAQGPNTNNASLGAMGKSPL 122
Qy 60 -----ATIIYITLPCSHFGSTPP 77
Db 123 NQGDSTPNLPKQAASTSGTTPP 145

RESULT 12
us-08-961-739-2
Sequence 2, Application US/08961739A
Patent No. 6063583
GENERAL INFORMATION:
APPLICANT: Montminy, Marc R.
TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
FILE REFERENCE: SALK1650-1
CURRENT APPLICATION NUMBER: US/08/961.739A
CURRENT FILING DATE: 1997-10-31
EARLIER APPLICATION NUMBER: US 194,468
EARLIER FILING DATE: 1994-02-10
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 2441
TYPE: PRT
ORGANISM: Mus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(2441)
OTHER INFORMATION: Xaa = Any Amino Acid
US-08-961-739-2

Query Match 12.2%; Score 61; DB 3; Length 2441;
Best Local Similarity 25.3%; Pred. No. 40;
Matches 21; Conservative 14; Mismatches 32; Indels 16; Gaps 3;
Qy 6 OLPNMVGXGTG--VNPPVGVVNEGRIVGIGAHLRKGDKHAQVQALDMAQXNAEG---- 59
Db 68 QLSLLRGSGSSINPGIGNVSASSPVQOGLG-----GQAQGPNTNNASLGAMGKSPL 122
Qy 60 -----ATIIYITLPCSHFGSTPP 77
Db 123 NQGDSTPNLPKQAASTSGTTPP 145

```
RESULT 13
US-08-912-227-2
; Sequence 2, Application US/08912227
; Patent No. 598171
; GENERAL INFORMATION:
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ni, Jian
; APPLICANT: Rosen, Craig A.
; TITLE OF INVENTION: Human Endokine Alpha
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,227
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,058
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0470001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 169 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-912-227-2

Query Match 12.1%; Score 60.5; DB 2; Length 169;
Best Local Similarity 43.8%; Pred. No. 1.2;
Matches 14; Conservative 3; Mismatches 4; Indels 11; Gaps 2;

QY 67 EPC-SHFG-----STPPCVNKIIDCKI 87
||| : || | ||||| : | :
Db 48 EPCMAKEGLPSKQWQMASSEPCVKNVSDKWL 79

RESULT 14
US-08-997-080-55
; Sequence 55, Application US/08997080
; Patent No. 598524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
```

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-080-55

Query Match 11.6%; Score 58; DB 2; Length 145;
Best Local Similarity 28.2%; Pred. No. 2.2;
Matches 22; Conservative 8; Mismatches 22; Indels 26; Gaps 4;

QY 19 PPVGAVVNEGRI-----VCIGAHLRK-GDKHAEVQALDMAQX 55
| | | | | : | | | | | : | | | | | : | : | :
Db 58 PAVDAVVISNDHYDHLIDITVALAHTQRAFPVPLGIGAHLRKGVPPEARIVELDMHEA 117

QY 56 N-AEGATYITLEPCSHF 72
: | : | : | : | : | :
Db 118 HRIDDLTLVCT--PARHF 133

RESULT 15
US-08-997-362-55
; Sequence 55, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
```

;
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleeth, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-997-362-55

Query Match 11.6%; Score 58; DB 2; Length 145;
Best Local Similarity 28.2%; Pred. No. 2.2;
Matches 22; Conservative 8; Mismatches 22; Indels 26; Gaps 4;
Qy 19 PPVGAVVYVNEGRI-----VGIGAHLRK-GDKHAEVQALDMAQX 55
| | | | | : | | | | | : | | :
Db 58 PAVDAVVISNDHYDHLDTITVALAHTORAPFVVPLGIGAHLRKKGVPPEARIVELDWHEA 117
Qy 56 N-AEGATIIYITLPCSHF 72
: : | : | | |
Db 118 HRIDDLTLVCT--PARHF 133

Search completed: November 14, 2000, 10:07:47
Job time: 65 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:08:13 ; Search time 24.32 Seconds
(without alignments)
258.329 Million cell updates/sec

Title: US-08-978-456-4

Perfect score: 498

Sequence: 1 MDYAIQLPNWOGXTGVNPP.....NKIIDCKIAXVVLXNKRQFR 99

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 182106

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_65: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	52.4	361	2 P01000	riboflavin-specific
2	249	50.0	356	2 C70313	riboflavin-specific
3	229	46.0	367	2 S26201	riboflavin-specific
4	216.5	43.5	339	2 F70901	probable riboflavin
5	214	43.0	369	2 B81039	riboflavin-specific
6	214	43.0	372	2 H64103	riboflavin-specific
7	213	42.8	357	2 G75552	riboflavin-specific
8	213	42.8	363	2 T10638	hypothetical prote
9	211	42.4	369	2 D81984	probable diamino
10	210.5	42.3	348	2 G72207	riboflavin-specific
11	196	39.4	368	2 S74377	riboflavin biosynt
12	193	38.8	376	2 G72026	riboflavin-specific
13	191	38.4	375	2 E71479	probable riboflavi
14	186	37.3	396	2 D81742	riboflavin-specific
15	125	25.1	336	2 D81258	probable riboflavi
16	118.5	23.8	148	2 D81879	probable cytosine
17	117.5	23.6	161	1 S11690	conserved hypothet
18	115.5	23.2	239	2 H81141	cytidine and deoxy
19	113.5	22.8	405	2 T41360	hypothetical prote
20	107	21.5	164	1 S74803	hypothetical prote
21	107	21.5	173	2 C64161	hypothetical prote
22	105.5	21.2	178	1 F65033	hypothetical 20.0
23	105	21.1	600	2 F75424	probable cell cycl
24	101	20.3	145	2 F75250	probable deoxycyti
25	99	19.9	599	2 T12994	riboflavin biosynt
26	96	19.3	151	2 G70377	conserved hypothet
27	92.5	18.6	201	2 B72380	hypothetical prote
28	88	17.7	188	1 DUBPC2	dCMP deaminase (EC
29	88	17.7	193	1 DUBPT4	dCMP deaminase (EC

ALIGNMENTS

RESULT 1

P01000

riboflavin-specific deaminase ribG - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 04-Dec-1992 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000

C/Accession: S45543; P01000; E69692

R:Sorokin, A.; Zumstein, E.; Azevedo, V.; Ehrlich, S.D.; Serror, P.

submitted to the EMBL Data Library, November 1993

A:Reference number: S45533

A:Accession: S45543

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-361 <SOR>

A:Cross-references: EMBL:L09228; NID:g410114; PIDN:AAA67481.1; PID:g410125

R:Mironov, V.N.; Perumov, D.A.; Krayev, A.S.; Stepanov, A.I.; Skryabin, K.G.

Mol. Biol. (Mosk.) 24, 256-261, 1990

A>Title: Unusual structure of Bacillus subtilis rib-operon regulatory region.

A:Reference number: P01000; MUID:90271920

A:Accession: P01000

A:Molecule type: DNA

A:Residues: 1-12 <MIR>

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Broutillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl

A:Authors: Schleich, S.; Schroeter, R.; Scroffone, F.; Sigiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili

A:Reference number: A69580; MUID:98044033

A:Accession: E69592

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-361 <KUN>

A:Cross-references: GB:299116; GB:AL009126; NID:g2634723; PIDN:CAB14260.1; PID:g26347

A:Experimental source: strain 168

C:Genetics:

A:Gene: ribG

C:Superfamily: Chlamydia pneumoniae riboflavin-specific deaminase

Query Match

52.4%; Score 261; DB 2; Length 361;

Best Local Similarity 56.7%; Pred. No. 5.4e-22;

Matches 51; Conservative 12; Mismatches 27; Indels 0; Gaps 0;

C:Function:
A:Description: converts 2,5-diamino-6-(ribosylamino)-4(3H)-pyrimidinone 5'-phosphate
to riboflavin biosynthesis
C:Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase
C:Keywords: hydrolase; riboflavin blosynthesis; zinc
F:50,75,84/Binding site: zinc (His, Cys, Cys) #status predicted

Query Match 46.0%; Score 229; DB 2; Length 367;
Best Local Similarity 51.6%; Pred. No. 2.3e-18;
Matches 47; Conservative 11; Mismatches 33; Indels 0; Gaps 0;

QY 1 MDVAIOLPNMVGXTCGNPPVGVAVVNVEGRIVGIGAHLRKGDKHAEOALDMAOXNASEGA 60
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 7 MARALKLAQRGRTPHPNPVNGCVIKDGEIVGEGYHORAGEPHAHVHALRMAGEKAKGA 66
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 61 TVITILEPCSHFGSTPPCCVKNIIDCKIAVV 91
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 67 TAYVTLEPCHSGHGRTPPCCDALIAAGVARV 97
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 4
F70901
probable riboflavin-specific deaminase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: F70901
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
J.; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
me sequence
A:Reference number: A70500; MUID:98295987
A:Accession: F70901
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-339 <COL>
A:Cross-references: GB:A123456; GB:AL123456; NID:g3256012; PID:CA02188.1; PID:g15429
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: ribG
C:Superfamily: Chlamydomophila pneumoniae riboflavin-specific deaminase

Query Match 43.5%; Score 216.5; DB 2; Length 339;
Best Local Similarity 48.9%; Pred. No. 5.6e-17;
Matches 45; Conservative 13; Mismatches 33; Indels 1; Gaps 1;

QY 1 MDVAIOLPNMVGXTCGNPPVGVAVVN-EGRIVGIGAHLRKGDKHAEOALDMAOXNAEG 59
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 13 MGIAIEHSYQVKGTTYPKPEVCATIVDPNGRIVGAGGTPEGACDHAEVVALRRAGCLAAG 72
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 60 ATIVITLEPCSIFGSTPPCCVKNIIDCKIAVV 91
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 73 AIWWTMEPCNHYGKTPPCVNALIEARVGTVV 104
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 5
B81039
riboflavin-specific deaminase NMBl817 [imported] - Neisseria meningitidis (group B st
rain 49226)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: B81039
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.;
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
et al.
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
The draft genome
A:Reference number: AB1000; MUID:20175755
A:Accession: B81039
A>Status: preliminary

RESULT **15**

D81258
probable riboflavin-specific deaminase (EC 3.5.4.-) Cj1622 [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C:date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C:Accession: D81258
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrett, T.; Brown, L.M.; Clark, F.; Clifton, H.; Davis, A.P.; De Jong, P.; et al., 2000
Nature 403, 665-668, 2000

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 14, 2000, 10:10:27 ; Search time 16.17 Seconds
(without alignments)
195.600 Million cell updates/sec

Title: US-08-978-456-4

Perfect score: 498

Sequence: 1 MDYAIQLPNMVGXTGVNPP.....NKIIDCKTAXVVLXNXRQFR 99

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 87993

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	261	52.4	361	1 RIBD_BACSU	P17618 b riboflavi
2	249	50.0	356	1 RIBD_AQUAE	O66534 a riboflavi
3	243	48.8	371	1 RIBD_BACAM	P70814 b riboflavi
4	229	46.0	367	1 RIBD_ECOLI	P25539 e riboflavi
5	216.5	43.5	339	1 RIBD_MYCTU	P71677 m riboflavi
6	214	43.0	372	1 RIBD_HAEIN	P44326 h riboflavi
7	196	39.4	368	1 RIBD_SYNY3	O55158 s riboflavi
8	195.5	39.3	337	1 RIBD_CORAM	O24750 c riboflavi
9	193	38.8	376	1 RIBD_CHLPN	Q92735 c riboflavi
10	180.5	36.2	376	1 RIBD_ACTPL	P50853 a riboflavi
11	117.5	23.6	161	1 YAAJ_BACSU	P21335 bacillus su
12	107	21.5	173	1 YFHC_HAEIN	P44931 haemophilus
13	105.5	21.2	178	1 YFHC_ECOLI	P30134 escherichia
14	96	19.3	151	1 Y903_AQUAE	O67050 aquifex aeo
15	88.5	17.8	147	1 YLXG_VIBFI	P33968 vibrio fisc
16	88	17.7	188	1 DCTD_BPT4	P00814 bacterioph
17	88	17.7	193	1 DCTD_HUMAN	P16006 bacterioph
18	83.5	16.8	178	1 DCTD_HUMAN	P32321 homo sapien
19	78	15.7	591	1 RIB2_YEAST	Q12362 saccharomyc
20	76	15.3	158	1 FCY1_YEAST	Q12178 saccharomyc
21	71.5	14.4	148	1 Y831_RICPR	Q22000 mycobacteri
22	70.5	14.2	128	1 DCTD_BPMDE	O22000 mycobacteri
23	69.5	14.0	197	1 DCTD_CAEEL	P30648 caenorhabdi
24	68	13.7	203	1 DCTD_DROME	O9vwa2 drosophila
25	66	13.3	568	1 URE2_HELHE	P42823 helicobacte
26	65.5	13.2	312	1 DCTD_YEAST	P06773 saccharomyc
27	64.5	13.0	150	1 FCAL_CANAL	P78594 candida alb
28	64.5	13.0	865	1 SECA_HELPY	O25475 helicobacte
29	63.5	12.8	211	1 HFA5_HAEIN	P45990 haemophilus
30	62.5	12.6	186	1 YD66_AQUAE	O67378 aquifex aeo
31	62	12.4	589	1 URE2_HELPY	P14917 helicobacte
32	61	12.2	352	1 RLA0_HALCU	P17006 halobacteri
33	61	12.2	352	1 RLA0_HALHA	P13553 halobacteri

RESULT 1

ID	RIBD_BACSU	STANDARD;	PRT;	361 AA.
AC	P17618;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD {INCLUDES: DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26) DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5- PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)}.			
GN	RIBD OR RIBG.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;			
OC	Bacillus/Staphylococcus group; Bacillus.			
[1]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / MARBURG;			
RX	MEDLINE; 95020538.			
RA	Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;			
RT	"The organization of the Bacillus subtilis 168 chromosome region between the spoVA and serA genetic loci, based on sequence data.";			
RL	Mol. Microbiol. 10:385-395(1993).			
[2]				
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168 / SHGW;			
RA	Mironov V.N.;			
[3]	Thesis (1989), USSR Academy of Sciences, Russia.			
RP	SEQUENCE OF 1-12 FROM N.A.			
RC	STRAIN=168 / SHGW;			
RX	MEDLINE; 90271920.			
RA	Mironov V.N., Perumov D.A., Kraev A.S., Stepanov A.I., Skryabin K.G.;			
RT	"Unusual structure of the regulatory region of the riboflavin biosynthesis operon in Bacillus subtilis.";			
RL	Mol. Biol. (Mosk) 24:256-261(1990).			
[4]				
RN	CHARACTERIZATION.			
RP	MEDLINE; 97221604.			
RA	Richter G., Fischer M., Krieger C., Eberhardt S., Luttgen H., Gerstenschlager I., Bacher A.;			
RT	"Biosynthesis of riboflavin: Characterization of the bifunctional deaminase-reductase of Escherichia coli and Bacillus subtilis.";			
RL	J. Bacteriol. 179:2022-2028(1997).			
CC	-1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)- 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE PYRIMIDINEDIONE 5'-PHOSPHATE.			
CC	-1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5- PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5- PHOSPHORIBOSYLAMINO)URACIL + NH(3).			
CC	-1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITV(LAMINO)URACIL + NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.			
CC	-1- COFACTOR: ZINC (BY SIMILARITY).			
CC	-1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.			
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES FAMILY.			


```

mutation." ;
RT Mol. Gen. Genet. 234:429-432(1992) .
RL [2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE: 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12." ;
RT Science 277:1453-1474(1997) .
RL [3]
RN
RP SEQUENCE FROM N.A.
RC
RX
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RA Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RL [4]
RN
RP CHARACTERIZATION.
RC MEDLINE: 97221604.
RX
RA Richter G., Fischer M., Krieger C., Eberhardt S., Luttgen H.,
RA Gerstenschlaeger I., Bacher A.;
RA "Biosynthesis of riboflavin: characterization of the bifunctional
RT deaminase-reductase of Escherichia coli and Bacillus subtilis." ;
RT J. Bacteriol. 179:2022-2028(1997) .
RC
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC
CC -2- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2O) = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3) .
CC
CC -3- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC
CC -4- COFACTOR: ZINC (BY SIMILARITY) .
CC
CC -5- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC
CC -6- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC
CC -7- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC EMBL: X64395; CAA45735.1; -
CC DR EMBL: AE000148; AAC73517.1; -
CC DR EMBL: U82664; AAB40170.1; -
CC DR PIR: S26201; S26201.
CC DR ECOGENE: EG11321; RIBD.
CC DR INTERPRO: IPR002125; -
CC DR INTERPRO: IPR002734; -
CC DR PFAM: PF01872; RIBD_C; 1.
CC DR PFAM: PF00383; dCMP_cyt_deam; 1.
CC DR PROSITE: P500903; CYP_DCMP_DEAMINASES; 1.
CC DR Riboflavin biosynthesis; Hydroxylase; Zinc; Oxidoreductase; NADP;
CC Multifunctional enzyme.
CC
CC DOMAIN 1 145 DEAMINASE.
CC FT DOMAIN 146 367 REDUCTASE.
CC FT METAL 50 50 ZINC (BY SIMILARITY) .
CC FT METAL 75 75 ZINC (BY SIMILARITY) .
CC FT METAL 84 84 ZINC (BY SIMILARITY) .
CC SEQUENCE 367 AA; 40338 MW; B19CEFA74D48D14D CRC64;
CC SQ

```

FT	METAL	57	57	ZINC (BY SIMILARITY).
FT	METAL	82	82	ZINC (BY SIMILARITY).
FT	METAL	91	91	ZINC (BY SIMILARITY).
SQ	SEQUENCE	339 AA;	33366 MW;	1DC0A1B8E7EC0B4 CRC64;

Query Match 43.5%: Score 216.5: DB 1: Length 339.

	437	Conservative	13	Mismatches	33	Indels	1	Gaps	1
QY	1	MDYAQLPNMQGXTGVNPPVGAVVYN-EGRIVIGIAHLRKGDKHAEEVQALDMAQXNAEG	59						
	:	:	:	:	:	:	:	:	:

Qy	60	ATVITTEPCSHFGSTPPCVNKIIDCKIAVW	91
		: : : : : :	
Db	73	AIVVTEPCNHYGKTTPCVNALIEARVGVV	104

RESULT 6
RIRD HAFIN

AC	01-NOV-1995	(Rel. 32, Created)
DT	01-NOV-1995	(Rel. 32, Last sequence update)
DT	01-NOV-1995	(Rel. 32, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOTIADIN BIOSYNTHESIS PROTEIN RIBD (INCLUDES:

DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)

DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-

DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].

GN RIBD OR RIBG OR HI0944;
QS Haemophilus influenzae

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae:

OC Haemophilus.

RN [1]

RP SEQUENCE FROM N.A.
PC SMPATN=PD / KJ20.

RX MEDLINE: 95350630.

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

RA Scott J.D., Shirley R., Liu L.-I., Glodex A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Springs T., Hedblom E., Cotton M.D.

RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venten + C

RT "Whole-genome random sequencing and assembly of *Haemophilus*
venereus

RT influenzae Rd." ;

RL Science 269:496-512(1995).

-1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE

CC 3'-PHOSPHATE INTO 3-AMINO-8-(RIBOSYLAMINO)-2,4-(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE

CC -I- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-

CC PHOSPHORIBOSYLAMINO) PYRIMIDINE + H(2)O = 5-AMINO-6-(5-

CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).

CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +

CC NADP(+) = 3-AMINO-6-(3-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY)

CC -1- PATHWAY: SECOND AND THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND

CC DEOXYCYTIDYLATE DEAMINASES FAMILY.

CC -[SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY

[illegible]

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is published in a peer-reviewed journal.

CC use by non profit institutions as long as its content is in no way modified and this statement is not removed Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

22

```
DR EMBL: U32775; AAC22598.1; -.
DR TIGR: HI0944; -.
DR INTERPRO: IPR002125; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF01872; RIBD_C; 1.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; 1.
KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 150 DEAMINASE.
FT DOMAIN 151 372 REDUCTASE.
FT METAL 55 55 ZINC (BY SIMILARITY).
FT METAL 80 80 ZINC (BY SIMILARITY).
FT METAL 89 89 ZINC (BY SIMILARITY).
FT METAL 89 89 ZINC (BY SIMILARITY).
SQ SEQUENCE 372 AA; 40931 MW; C7504EDDB8B86F7D CRC64;

Query Match 43.0%; Score 214; DB 1; Length 372;
Best Local Similarity 47.3%; Pred. No. 7.6e-17;
Matches 43; Conservative 11; Mismatches 37; Indels 0; Gaps 0;

QY 1 MDYAIQLPNMVOGTYGPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 12 MQRALDLAARGQYTTTPNPSCVGLVNGEIVGEGFHFKAGQPHAEKVAQAAGENAKGA 71
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 61 TIYTITPCSHFGSTPPCVNKNIIDCKIAXV 91
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 72 TAYVLEPCAHYGTTPCALGLIAGVVKVI 102
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 7
RIBD_SYNY3
ID RIBD_SYNY3 STANDARD; PRT; 368 AA.
AC O55158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR SLR0066.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
[1]
RN SUGIURA M., TANAKA A., SATO S., KOTANI H., SAZUKA T., MIYAJIMA N.,
RA Sugiyura M., Tanaka S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5'-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: D64001; BAA10295.1; -.
DR INTERPRO: IPR002125; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF01872; RIBD_C; 1.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; 1.
KW Riboflavin biosynthesis; Hydrolase; Zinc; Oxidoreductase; NADP;
KW Multifunctional enzyme.
FT DOMAIN 1 146 DEAMINASE.
FT DOMAIN 147 368 REDUCTASE.
FT METAL 51 51 ZINC (BY SIMILARITY).
FT METAL 76 76 ZINC (BY SIMILARITY).
FT METAL 85 85 ZINC (BY SIMILARITY).
SQ SEQUENCE 368 AA; 39995 MW; BCDA8ED0916B50BC CRC64;

Query Match 39.4%; Score 196; DB 1; Length 368;
Best Local Similarity 45.7%; Pred. No. 8e-15;
Matches 42; Conservative 10; Mismatches 40; Indels 0; Gaps 0;

QY 1 MDYAIQLPNMVOGTYGPPVAVVNEGRIVGIGAHLRKGDHAEVQALDMAQXNAEGA 60
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 8 MRRCLTLAKTAIGKATPNPLVGSIVGQGFHPQAGQPHGEIPALWEAGDRAKGA 67
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 61 TIYTITPCSHFGSTPPCVNKNIIDCKIAXV 92
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 68 TLYVNLPECNHQRTPPCTEALIQAGIAKVVV 99
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 8
RIBD_CORAM
ID RIBD_CORAM STANDARD; PRT; 337 AA.
AC O24750;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBD [INCLUDES:
DE DIAMINOHYDROXYPHOSPHORIBOSYLAMINOPYRIMIDINE DEAMINASE (EC 3.5.4.26)
DE (RIBOFLAVIN-SPECIFIC DEAMINASE); 5-AMINO-6-(5-
DE PHOSPHORIBOSYLAMINO)URACIL REDUCTASE (EC 1.1.1.193) (HTP REDUCTASE)].
GN RIBD OR RIBG.
OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
OC Corynebacterium.
[1]
RN SEQUENCE FROM N.A.
RA Koizumi S., Yonetani Y., Teshiba S.;
RT "Process for producing riboflavin.";
RT Patent number US5589359, 31-DEC-1996.
CC -1- FUNCTION: CONVERTS 2,5-DIAMINO-6-(RIBOSYLAMINO)-4(3H)-PYRIMIDINONE
CC 5'-PHOSPHATE INTO 5'-AMINO-6-(RIBOSYLAMINO)-2,4(1H,3H)-
CC PYRIMIDINEDIONE 5'-PHOSPHATE.
CC -1- CATALYTIC ACTIVITY: 2,5-DIAMINO-6-HYDROXY-4-(5-
CC PHOSPHORIBOSYLAMINO)PYRIMIDINE + H(2)O = 5-AMINO-6-(5-
CC PHOSPHORIBOSYLAMINO)URACIL + NH(3).
CC -1- CATALYTIC ACTIVITY: 5-AMINO-6-(5-PHOSPHORIBITYLAMINO)URACIL +
CC NADP(+) = 5-AMINO-6-(5-PHOSPHORIBOSYLAMINO)URACIL + NADPH.
CC -1- COFACTOR: ZINC (BY SIMILARITY).
CC -1- PATHWAY: SECOND AN THIRD STEPS OF RIBOFLAVIN BIOSYNTHESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE CYTIDINE AND
CC DEOXYCYTIDYLATE DEAMINASES FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE HTP
CC REDUCTASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```


CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR	EMBL; D26185; BAA05254.1; -				
DR	EMBL; Z99104; CAB11794.1; -				
DR	PIR; S11690; S11690.				
DR	SUBTILIST; BG10082; YAAJ.				
DR	INTERPRO; IPR002125; -				
DR	PFAM; PF00383; dCMP_cyt_deam; 1.				
DR	PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.				
KW	Hypothetical protein; Hydrolase; Zinc.				
FT	METAL 53 53	ZINC (BY SIMILARITY).			
FT	METAL 83 83	ZINC (BY SIMILARITY).			
FT	METAL 86 86	ZINC (BY SIMILARITY).			
SQ	SEQUENCE 161 AA; 17751 MW; 3256GF31DF6610FBB CRC64;				

RESULT	ID	YFHC_HAEIN	STANDARD;	PRT;	173 AA.
AC	P44931;				
DT	01-NOV-1995	(Rel. 32, Created)			
DT	01-NOV-1995	(Rel. 32, Last sequence update)			
DT	01-NOV-1995	(Rel. 32, Last annotation update)			
DE	HYPOTHETICAL PROTEIN HI0906.				
GN	HI0906.				
OS	Haemophilus influenzae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;				
OC	Haemophilus.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=RD / KW20.				
RX	MEDLINE; 95350630.				
RA	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,				
RA	Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,				
RA	McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,				
RA	Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,				
RA	Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,				
RA	Uitterlind T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,				
RA	Fine L.B., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,				
RA	Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,				
RA	Venter J.C.;				
RT	"whole-genome random sequencing and assembly of Haemophilus				

RL Science 269:496-512(1995).
 CC -|- COFACTOR: ZINC (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
 CC FAMILY. STRONG, TO E.COLI AND B.SUBTILIS RIBG. STRONG, TO E.COLI
 CC YFHC.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

```
DR EMBL; U32772; AAC22565.1; -.
DR TIGR; H10906; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 61 61 ZINC (BY SIMILARITY).
FT METAL 91 91 ZINC (BY SIMILARITY).
FT METAL 94 94 ZINC (BY SIMILARITY).
SQ SEQUENCE 173 AA; 19359 MW; 785B2CE532FB72DE CRC64;

Query Match 21.5%; Score 107; DB 1; Length 173;
Best Local Similarity 32.7%; Pred. No. 3.8e-05;
Matches 33; Conservative 20; Mismatches 30; Indels 18; Gaps 6;

QY 1 MDVAIQLPNVVOGTGYNPPGAVVNEGR-IVGIGAHLR--KCD--KHAEOVQALDMAQX 55
| | | | | : : | | | | | : : | | | : : | | | : |
15 MRYALELADRAEALGETI--PVGAVLYDDARNIIGEGWNLIVSDPTTAHAEIILNRGAK 72

QY 56 NAE-----GATVITLPCSHFGSTPPCVNKLIDCKIAXVV 91
| : : : | | | | | : | : : | |
73 NIQYRLNLTSLVYLEPCT-----MCAGAILHSRIKRLV 107

Db 73 NIQYRLNLTSLVYLEPCT-----MCAGAILHSRIKRLV 107

RESULT 13
YFHC_ECOLI
ID YFHC_ECOLI STANDARD; PRT; 178 AA.
AC P30134;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 20.0 KDA PROTEIN IN PURL-DPJ INTERGENIC REGION (ORF178).
GN YFHC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NWL37;
RX MEDLINE; 92292954.
RA Poulsen L.K., Larsen N.W., Molin S., Andersson P.;
RT "Analysis of an Escherichia coli mutant strain resistant to the cell-
RT killing function encoded by the gef gene family.";
RL Mol. Microbiol. 6:895-905(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Nashimoto H., Salto N.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE; 97426617.
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1433-1474(1997).
CC -!- FUNCTION: MUTATION IN THIS PROTEIN MAKES E. COLI RESISTANT TO THE
CC TOXIC PROTEINS ENCODED BY THE GEF GENE FAMILY.
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
CC FAMILY. STRONG. TO E. COLI AND B. SUBTILIS RIBG. STRONG, TO
CC H. INFLUENZAE HI0906.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; A5000713; AAC07025.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 52 52 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).

or send an email to license@isb-sib.ch).
-----
CC EMBL; X72336; CAA51064.1; -.
DR EMBL; D64044; BAA10909.1; -.
DR EMBL; U36841; AAA79821.1; -.
DR EMBL; AE000342; AAC75612.1; -.
DR PIR; S20974; S20974.
DR ECOGENE; EG11372; YFHC.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 68 68 ZINC (BY SIMILARITY).
FT METAL 98 98 ZINC (BY SIMILARITY).
FT METAL 101 101 ZINC (BY SIMILARITY).
FT MUTAGEN 64 64 D->E: RESISTANCE TO THE CELL-KILLING
FUNCTION ENCODED BY THE GEF GENE FAMILY.
SQ SEQUENCE 178 AA; 20026 MW; 80B2E3B5FD61AA8A CRC64;

Query Match 21.2%; Score 105.5; DB 1; Length 178;
Best Local Similarity 36.1%; Pred. No. 5.7e-05;
Matches 30; Conservative 9; Mismatches 25; Indels 19; Gaps 4;

QY 20 PVGAVVNEGRIVG-----IGAHLRKDKHAEVQALD-----MAQXNAEGATYYITLEP 68
| | | | | : : : | | | | | : : : | | | | |
40 PVGAVLVHNNRVIGEGWNRPIGRH--DPTAHAEIMALRQGLVMQNYRLIDATLYVTLEP 97

QY 69 CSHFGSTPPCVNKLIDCKIAXVV 91
| : : : | | : | : |
98 C-----VMCAGAMIHSRIGRVV 114

Db 98 C-----VMCAGAMIHSRIGRVV 114

RESULT 14
Y903_AQUAE
ID Y903_AQUAE STANDARD; PRT; 151 AA.
AC O67050;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN AQ_903.
GN AQ_903.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE; 98196666.
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; A5000713; AAC07025.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 52 52 ZINC (BY SIMILARITY).
FT METAL 82 82 ZINC (BY SIMILARITY).
```

Job time: 177 sec

```
FT METAL      85      85      ZINC (BY SIMILARITY).
SQ SEQUENCE 151 AA; 17246 MW; 5543DC00B9846D25 CRC64;

Query Match      19.3%; Score 96; DB 1; Length 151;
Best Local Similarity 35.4%; Pred. No. 0.00056;
Matches 29; Conservative 10; Mismatches 19; Indels 24; Gaps 5;

QY 20 PVGAVVNEGRIVGIGAH-----LRKGDKHAEVQALDMA--QXNA---EGATYITILEPC 69
   ||||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24 PVGAIVKEGEIIS-KAHNSVEELKDPTAHAEMLAIKEACRLNTKYLEGCGLYVTLEPC 82

QY 70 SHFGSTPPCVKNKIIDCKIAXV 91
   | | | | |
Db 83 -----IMCSYALVL 91
```

```
RESULT 15
YILXG_VIBFI
ID YLXG_VIBFI STANDARD; PRT; 147 AA.
AC P33968;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE HYPOTHETICAL 16.6 KDA PROTEIN IN LUXG 3 REGION.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 7744;
RX MEDLINE; 93320098.
RA Lee C.Y., Zittner R.B., Miyamoto C.M., Meighen E.A.;
RT "The gene convergent to luxG in Vibrio fischeri codes for a protein
RL Biochim. Biophys. Acta 1143:337-339(1993).
CC -!- COFACTOR: ZINC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTIDINE AND DEOXYCYTIDYLATE DEAMINASES
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X70289; CAA49769.1; -.
DR INTERPRO; IPR002125; -.
DR PFAM; PF00383; dCMP_cyt_deam; 1.
DR PROSITE; PS00903; CYT_DCMP_DEAMINASES; 1.
KW Hypothetical protein; Hydrolase; Zinc.
FT METAL 67 67 ZINC (BY SIMILARITY).
FT METAL 92 92 ZINC (BY SIMILARITY).
FT METAL 95 95 ZINC (BY SIMILARITY).
SQ SEQUENCE 147 AA; 16622 MW; 174FDA8DC1458D49 CRC64;
```

```
Query Match      17.8%; Score 88.5; DB 1; Length 147;
Best Local Similarity 30.3%; Pred. No. 0.0038;
Matches 27; Conservative 10; Mismatches 25; Indels 27; Gaps 3;

QY 21 VGAVVNEGRIVGIG-----AHLRGDKHAEVQALDMAQXNAEGAT 61
   ||||:| | | | | | | | | | | | | | | | | | | | | | | |
Db 27 VGAVITKINRIVSGFNGYPHCVSDSADTDREIKYLK--TLHAENAILFAKRDLEGCD 84

QY 62 IYITLEPCSHFGSTPPCVKNKIIDCKIAXV 90
   | : | | | | | | | | | | | | | | | | | | | | | | |
Db 85 IWTHTFPC-----PNCAAKIIQTGISKV 107
```

THIS PAGE BLANK (USPTO)


```
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE.
GN DR0153.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL;
RX MEDLINE; 20036896.
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome Sequence of the Radioreistant Bacterium Deinococcus
RT radiodurans RL.";
RL Science 286:1571-1577(1999).
RL EMBL: AE001878; AAF09742.1; -.
DR TIGR: DR0153; -.
DR INTERPRO: IPR002125; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PFAM: PF01872; RibD_C; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; 1.
SQ SEQUENCE 357 AA; 37226 MW; ADBFE65C10626007 CRC64;

Query Match 42.8%; Score 213; DB 2; Length 357;
Best Local Similarity 52.4%; Pred. No. 2.8e-16;
Matches 44; Conservative 9; Mismatches 27; Indels 4; Gaps 1;

QY 13 GXTGVNPPVGVVNEG----RVIGTGAHLRKDGKHAQVQALDMAQXNAEGAIYITLPE 68
Db 11 ||||| : : : : : ||||| : : : : : ||||| : : : : : |||||
QY 69 CSHFGSTPPCVNKIKDCKIAKXVVL 92
Db 111 | |||| : : : : : ||| : : : : :
QY 84 CSHDGRTPPCADALIAGVARVV 107
Db 111 | |||| : : : : : ||| : : : : :

RESULT 3
Q9SUB7 PRELIMINARY; PRT; 363 AA.
AC Q9SUB7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE HYPOTHETICAL 39.7 KDA PROTEIN.
GN T13K14.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC Brassicaceae; Arabidopsi.
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL080282; CAB45891.1; -.
DR INTERPRO: IPR002125; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
KW Hypothetical protein
SQ SEQUENCE 363 AA; 39729 MW; 1281410D8FE272F6 CRC64;
```

```
Query Match 42.8%; Score 213; DB 10; Length 363;
Best Local Similarity 46.7%; Pred. No. 2.9e-16;
Matches 43; Conservative 12; Mismatches 37; Indels 0; Gaps 0;

QY 1 MDVAIQLPNMVGXGTGVNPPVGVVNEGRIVGIGAHLRKDGKHAQVQALDMAQXNAEGA 60
Db 11 : : : : : ||||| : : : : : ||||| : : : : : |||||
QY 61 TIVITLPCSHFGSTPPCVNKIKDCKIAKXVVL 92
Db 11 : : ||||| : : ||||| : : ||| : : : : : |||
QY 75 TAYVLEPCNHYGRTPTCTEALINAKVRRVVI 106
Db 11 : : ||||| : : ||||| : : ||| : : : : : |||

RESULT 4
Q9REF6 PRELIMINARY; PRT; 363 AA.
AC Q9REF6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE / REDUCTASE (FRAGMENT).
GN RIBD.
OS Bartonella henselae (Rochalimaea henselae).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bartonellaceae; Bartonella.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOUSTON-1;
RA Bereswill S., Hinkelmann S., Kist M., Sander A.;
RT "Molecular analysis of riboflavin synthesis genes in bartonella
RT henselae and use of the ribC gene for differentiation of bartonella
RT species by PCR.";
RL J. Clin. Microbiol. 37:3159-3166(1999).
DR EMBL: AJ132928; CAB63090.1; -.
DR INTERPRO: IPR000886; -.
DR INTERPRO: IPR002125; -.
DR INTERPRO: IPR002734; -.
DR PFAM: PF00383; dCMP_cyt_deam; 1.
DR PFAM: PF01872; RibD_C; 1.
DR PROSITE: PS00014; ER_TARGET; UNKNOWN_1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 363 AA; 39524 MW; 131630AF56074F7D CRC64;

Query Match 42.5%; Score 211.5; DB 2; Length 363;
Best Local Similarity 47.9%; Pred. No. 4.3e-16;
Matches 45; Conservative 10; Mismatches 34; Indels 5; Gaps 1;

QY 4 AIQLPNMVGXTGVNPPVGVVNEG-----IVGIGAHLRKDGKHAQVQALDMAQXNAE 58
Db 11 : : ||||| : : ||||| : : ||| : : : : : |||
QY 2 AIRLAERHVGLTGENSVGTIIARNDENVCVYIVGVGTATQGRPHAEVQALOMAGSLAH 61
Db 111 | ||||| : : ||||| : : ||| : : : : : |||
QY 59 GATITLPCSHFGSTPPCVNKIKDCKIAKXVVL 92
Db 111 | ||||| : : ||||| : : ||| : : : : : |||
QY 62 GATAYVTLPCSHYGTSPCVNTLLKSGISRVI 95
Db 111 | ||||| : : ||||| : : ||| : : : : : |||

RESULT 5
Q9X2E8 PRELIMINARY; PRT; 348 AA.
AC Q9X2E8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE RIBOFLAVIN-SPECIFIC DEAMINASE.
GN TM1828.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB / DSM 3109;
```

[illegible]

DR	PFAM; PF00383; dCMP_cyt_deam; 1.
DR	PROSITE; PS00903; Cyt_DCMP_DEAMINASES; UNKNOWN_1.
KW	Hypothetical protein.
SQ	SEQUENCE 164 AA; 17985 MW; EE03A8329069459D CRC64;
Query Match 21.5%; Score 107; DB 2; Length 164;	
Best Local Similarity 34.1%; Pred. No. 0.00011;	
Matches 28; Conservative 13; Mismatches 25; Indels 16; Gaps	
Qy	20 PVGAVVNE-GRIVIGAHLRKGDK----HAEVALDMA-----QXNAEGATVIITLPC 69 : : : : :
Db	29 PVGAVVNAMGEIATCONRRQRDNPTAAEMLAIOACRRLGHWRNLECTLVTLPC 88 : : : : :
Qy	70 SHEGSTPCVNKIIDCKIAVV 91 : : : : :
Db	89 -----PMTGAILQARGLLV 104 : : : : :
RESULT 11	
Q9RKM1	PRELIMINARY; PRT; 376 AA.
AC Q9RKM1	
DT 01-MAY-2000	(TrEMBLrel. 13, Created)
DT 01-MAY-2000	(TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2000	(TrEMBLrel. 14, Last annotation update)
DE PUTATIVE BIFUNCTIONAL ENZYME DEAMINASE/REDUCTASE.	
GN SCD17.10.	
OS Streptomyces coelicolor.	
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;	
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.	
[1]	
RP SEQUENCE FROM N.A.	
RC STRAIN-A3(2);	
RA Brown S.P., Harris D.;	
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.	
[2]	
RP SEQUENCE FROM N.A.	
RC STRAIN-A3(2);	
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;	
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.	
[3]	
RP SEQUENCE FROM N.A.	
RC STRAIN-A3(2);	
RA Redenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J.,	
RA Kinashi H., Hopwood D.A.;	
RT "A set of ordered cosmids and a detailed genetic and physical map for	
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";	
RT Mol. Microbiol. 21:77-96(1996).	
DR EMBL; AL118515; CAB56387.1; -.	
DR INTERPRO; IPR002125; -.	
DR INTERPRO; IPR002734; -.	
DR PFAM; PF00383; dCMP_cyt_deam; 1.	
DR PFAM; PF01872; RibD_C; 1.	
SQ SEQUENCE 376 AA; 39583 MW; 3D3729BD61AFE6BA CRC64;	
Query Match 21.2%; Score 105.5; DB 2; Length 376;	
Best Local Similarity 40.2%; Pred. No. 0.00044;	
Matches 35; Conservative 8; Mismatches 29; Indels 15; Gaps	
Qy 19 PP-----VGAVVV-NEGRIVGICAHLRKKDGK---HAEVQAL---DMAQXNAEGATVIIT 65 : : : :	
Db 246 PPSDATSVGVVAADSGEALRG-HSGEDGPVVHAEALAKVDPPRLPGATVISS 300 : : : :	
Qy 66 LEPSCSHFGSTP-PCNVKIIDCKIAVV 91 : : : :	
Db 305 LEPCARRASRPAPCARLILDAGVRVV 331 : : : :	
RESULT 12	
Q9RV23	PRELIMINARY; PRT; 600 AA.
ID Q9RV23	

RL	Submitted (AUG-1999) to the EMBL/GenBank/DBSJ databases.	
DR	EMBL; AC011507.1; -	
DR	EMBL; AC008075; AAD49971.1; -	
DR	INTERPRO: IPR002125; -	
DR	PFAM: PF00383; dCMP_cyt_deam; 1.	
SO	SEQUENCE 1307 AA; 146563 MW; 8CDAD95ABAB24719 CRC64;	
Query Match 21.0%; Score 104.5; DB 10; Length 1307;		
Best Local Similarity 31.0%; Pred. No. 0.0025;		
Matches 27; Conservative 14; Mismatches 19; Indels 27; Gaps		
QY	20 PVGAVVWNEGRIVGIG-----AHLRKGDGKHAEOALDMAOXNAEGA-----TIIYI 64	
	: : : : :	:
Db	1131 PVGAVLVHDGKIITARGVIVVEELRSDTAHEMICI-----REGSKALRSWRLADTTLVY 1184	
QY	65 TLEPCSHFGSTPPCVNKIIDCKTAXVV 91	
	: : :	
Db	1185 TLEPC-----PMCAGAILQARVNTLV 1205	
RESULT 14		
Q9RR66	PRELIMINARY; PRT; 145 AA.	
ID	Q9RR66	
AC	Q9RR66	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT	01-JUN-2000 (TrEMBLrel. 14, Last annotation update)	
DE	DEOXYCYTIDYLATE DEAMINASE, PUTATIVE.	
GN	DR2631.	
OS	Deinococcus radiodurans.	
OC	Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.	
NC	[1]	
RP	SEQUENCE FROM N.A.	
RN	STRAIN-RL;	
RC	MEDLINE; 20036896.	
RX	RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,	
RA	Madson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,	
RA	Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,	
RA	Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,	
RA	Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,	
RA	Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,	
RA	Fraser C.M.;	
RT	"Genome Sequence of the Radioresistant Bacterium Deinococcus	
RT	radiodurans RL";	
RL	Science 286:1571-1577(1999).	
DR	EMBL; AE002092; AAF12167.1; -	
DR	TIGR; DR2631; -	
DR	INTERPRO: IPR002125; -	
DR	PFAM: PF00383; dCMP_cyt_deam; 1.	
DR	PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.	
SO	SEQUENCE 145 AA; 15547 MW; 110CA05197EE999A CRC64;	
Query Match 20.3%; Score 101; DB 2; Length 145;		
Best Local Similarity 34.1%; Pred. No. 0.00046;		
Matches 28; Conservative 11; Mismatches 21; Indels 22; Gaps		
QY	21 VGAVVWV-EGRIVGIGAHLR-----KGDKHAEVQALDMAOXNAEGATIIYI 64	
	: : : :	
Db	28 VGACILDRHRVGVGYNGRAAGEPNERESLAQGASGYTHAEVNALLAANWNGEGTTLVY 87	
QY	65 TLEPCSHFGSTPPCVNKIIDCK 86	
	: : :	
Db	88 THEPCS-----VCARLIVNSR 103	
RESULT 15		
Q9STY4	PRELIMINARY; PRT; 599 AA.	
ID	Q9STY4	
AC	Q9STY4	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	

